Stochastic Modelling of Coral-Algal Symbiosis on the Great Barrier Reef

Max Wurm

Thesis submitted for the degree of
Master of Philosophy
in
Applied Mathematics
at The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences

School of Mathematical Sciences

THE UNIVERSITY
of ADELAIDE

February 2020

In loving memory of my babcia,

Eve Wurm,
(March 1933 – January 2020).

You taught me resilience

and a bunch of other fancy words.

Contents

Abstract

The Great Barrier Reef is the largest coral reef system on earth, but is threatened by the phenomenon of coral bleaching. We create a stochastic fluid flow model for coral bleaching, with the aim of better understanding its underlying mechanisms. Analysing this model involves the inversion of Laplace-Stieltjes transforms, a process ripe with difficulties. A recently developed method for Laplace transform inversion, called the Concentrated Matrix Exponential method, is very effective in overcoming these difficulties. Proceeding our analysis, we explore the concept of Parisian ruin to improve the biological realism of the model, which inspires a novel modelling framework for the problem. This framework includes an explicit energy process for the coral, and takes into account the traits of different algal species. We find that under our model, corals can benefit from hosting the two most prevalent algal species on the Great Barrier Reef, as opposed to only one species.

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: ..... Date: .....

Acknowledgements

To my supervisors Professor Nigel Bean and Dr Giang Nguyen: an *immeasurable* thank-you for everything: your belief in me from the very start; your tireless committment to my project; your transcendent wisdom and knowledge of mathematics; your ability to motivate me when nothing works and the answers don't exist. The last two years have been truly amazing, and I owe that all to you two.

To Mum, Dad and Zoe: a *million* thank-you's for your unfaltering love and support throughout these past two years. I couldn't have made it through without you, and I wholeheartedly appreciate everything you do for me.

To Caitlin: thank you *so* much for the countless tea breaks and puzzle sessions, for listening to me complain over and over about my code not working, and for offering me chocolate every so often. I appreciate it so much.

To Sophie: a *sincere* thank-you for all the tea breaks and procrastination sessions which fueled this thesis. And always remember: buffalo buffalo buffalo buffalo buffalo buffalo buffalo buffalo.

To Angus and Dennis: *all* the thank-you's for your help and advice, for listening to me complain over and over about my CDFs not going to one, and, of course, for feeding my banh mi addiction.

To Han, Harry, Alex and Jess: an *immense* thank-you for keeping me sane, and enabling me to cling on to some semblance of a social life.

To Ellie and Rhaneela: an *infinite* thank-you for all the whiteboard sessions, hot chips, and for helping me live my life to the lim sup.

To Sean Connolly, Andrew Baird, and the friendly PhD students at James Cook University: a *Great (Barrier)* thank-you for making me feel welcome in Townsville, for the many helpful talks, and for enabling me to understand coral bleaching at a much deeper level than I could ever have hoped to achieve sitting at my desk in Adelaide, thousands of kilometres away from the Great Barrier Reef.

A *big* thank-you to ACEMS for sponsoring me on many occasions, including the conferences I was fortunate enough to present at, and my trip to Townsville.

Finally, to Zen Kitchen: a *delicious* thank-you for sustaining me throughout my five years at university. You will always have the best banh mi in Adelaide.

Chapter 1

Introduction

1.1 The quarrel in the coral

Coral reefs are among the most incredibly beautiful, complex and diverse ecosystems on earth, and have been around for many thousands of years. The corals that build these reefs are fascinating animals due to their symbiotic relationship with a type of algae colloquially known as *zooxanthellae*.

There are many coral reefs of various shapes, sizes and types all across the globe, but there is one that stands out above them all (particularly from the eyes of an Australian).

The Great Barrier Reef (GBR), located off the coast of North-East Australia, is the largest coral reef system on earth, so large that it is visible from space. Its 2,500 individual reefs span approximately 348,000km^2 and house over 400 species of coral, 1,500 species of fish, and 4,000 species of mollusc [21], making the reef a cornerstone of marine biodiversity and natural beauty.

Furthermore, the GBR is a significant economic asset to Australia, contributing more than A$1.2 billion per annum (p.a.) to the Australian economy via tourism (∼A$700 million p.a.), commercial fishing (∼A$250 million p.a.) and recreational boating/fishing (∼A$270 million p.a.) [29]. This economic importance extends globally, since coral reefs house approximately 10% of fish consumed by humans worldwide, despite covering less than half a percent of the ocean floor [58].

Currently, the biggest threat to the GBR, and other coral reefs worldwide, is *coral bleaching*, a phenomenon whereby coral loses its energy-producing symbiotic partner, and, as a result, becomes energy-starved. This starvation can lead to lack of growth [20], increased susceptibility to disease [17], and, if the bleaching is prolonged, death.

Coral bleaching can be triggered by a range of environmental factors, but widespread bleaching events have been linked directly to elevated sea-surface temperatures (SST) due to climate change [52]. The most recent and significant of these events, the 'Third Global Coral Bleaching Event', wiped out an estimated 29% of coral on the GBR in 2014 and 2015 [38].

Notwithstanding attempts to mitigate the effect of climate change, mass bleaching events are predicted to continue and increase in frequency in the next few decades [30], provoking the need for research in this area.

This thesis aims to understand the processes of coral bleaching using mathematical models. Existing models tend to focus on the large-scale dynamics of coral bleaching, rather than the small-scale dynamics; moreover, small-scale models often do not account for the various sources of randomness associated with coral bleaching, such as environmental conditions, ecological factors, and other unknown forces. By formulating a mathematical model which accounts for this randomness, we can understand coral bleaching on an entirely new level, and help mitigate the negative effects of coral bleaching in the future.

The models we consider here focus on the small-scale dynamics of coral-algal symbiosis, the important relationship between coral and zooxanthellae. Coral bleaching is a result of a coral expelling its zooxanthellae population, and so understanding this symbiosis is key to understanding coral bleaching.

To model this relationship, we use a *stochastic fluid model*, due to its flexibility, and ability to capture temporal behaviour. In this model, we consider the algae population as a continuous density which increases and decreases over time. Increasing density corresponds to population growth, whereas decreasing density corresponds to a bleaching event.

There are many species of zooxanthellae that can live within corals on the GBR. Each species has a different resistance to bleaching, growth rate, and ability to produce energy for the coral host, making the inter-species dynamics vital to understand coral bleaching. This is a key aspect of coral-algal symbiosis that we capture in our model.

We are also presented with an excellent opportunity to explore new types of stochastic fluid models, which is another focus of this work. While we begin with a relatively simple model, the need for more sophisticated modelling tools to understand coral bleaching provides an avenue for innovation. In later chapters, we explore novel extensions of stochastic fluid models, inspired by the desire to make our model more biologically realistic. In these models, mortality is defined by a *horizon*—a secondary process which runs in parallel to the fluid process.

1.2 Thesis summary

In Chapter 2, we lay the biological and mathematical foundation for the thesis. The biology behind coral bleaching is crucial to understanding existing models for coral bleaching, and developing new models. Following the biology of coral bleaching, we review current models for coral-algal symbiosis, and identify areas to improve upon this modelling. The latter half of the chapter is dedicated to the mathematics used to formulate our models for coral-algal symbiosis. We cover the Abate-Whitt framework, which is a routine for inverting Laplace (and Laplace-Stieltjes) transforms; we also cover the theory behind stochastic fluid flow models, which are the class of mathematical models that we use in our work.

The first model appears in Chapter 3, in which we state our key assumptions, specify the model mathematically, and specify a set of canonical parameters to use with the model. We derive the distribution of the *time to mortality*, which is an important measure for the survivability of the coral host, and reappears frequently in later chapters. It is important to note that obtaining the *distribution* provides far more information about the mortality process than a simpler measure such as the mean, and is possible to obtain only via a stochastic approach.

The time to mortality reappears immediately in Chapter 4, as the focal point. We obtain the distribution of the time to mortality by inverting its Laplace-Stieltjes transform. This inversion process is ripe with obstacles, so we investigate ways to overcome these issues. In particular, we explore a *time-shifting* method which allows for the inversion of functions with sharp jumps. We use this method to obtain the distribution of the time to mortality and compare the efficacy of various inversion methods.

In Chapter 5 we take a new path, and introduce a more sophisticated model for coral-algal symbiosis. This model allows for a more general definition of mortality, using the concept of *Parisian ruin*[1]. In this new model, mortality is defined as the time that a random horizon is reached. We formulate this model mathematically, and once again obtain the Laplace-Stieltjes transform of the time to mortality. Inverting this transform is inefficient with our current methods; as an example, we simulate the model using an Erlang distribution to explore the model in more detail. We show that under this model, we obtain the distribution of the time to mortality from the basic model in Chapter 3 in the limit.

An even more sophisticated model, called the *energy model*, is introduced in Chapter 6, inspired by the Parisian ruin model. This model includes an explicit

[1]Parisian ruin is a type of financial ruin which occurs when a quantity (e.g., the amount of money in a bank account) spends a certain amount of time below a given threshold (e.g., 0).

energy process for the coral host, and allows us to incorporate the advantages of different algal species, an element that was missing from previous models. We construct this model by defining a two-way horizon process and adding this to our fluid model from Chapter 3. We then introduce a level-dependent model, and from this model define the Laplace-Stieltjes transform of the time to mortality. The final part of this transform is very difficult to specify mathematically, so we are unable to obtain the distribution of the time to mortality via Laplace-Stieltjes transform inversion, but we can still simulate the energy model to obtain this distribution. Using simulation, we compare the energy model to simpler models which only allow for the coral to host a single algal species, and discuss our findings.

Finally, we conclude in Chapter 7 and discuss future directions for this research.

Reader beware: mathematical notation and abbreviations

Throughout this thesis, we use a wealth of abbreviations and mathematical notation. Although we strive to define every abbreviation the first time it is used, it may be difficult to remember the meaning of every acronym, symbol and decoration. We direct any confused readers to the glossary in Appendix A.

Chapter 2

Background

This chapter provides the necessary background information for the rest of the thesis. First, we take a closer look at the biology behind coral bleaching on the Great Barrier Reef (GBR). In addition, we briefly review previous models for coral bleaching, and motivate the need for a mechanistically driven stochastic model. Next, we cover the major mathematical concepts used throughout the thesis, split into two sections: the Abate-Whitt framework, and stochastic fluid flow models. The Abate-Whitt framework is a mathematical framework for the inversion of Laplace transforms, underpinning the numerical analysis in Chapters 4 and 5. We thoroughly explore it, and various methods within the framework. This thesis revolves around stochastic fluid flow models, and so we cover them extensively also. We construct a generic model from scratch, and derive important quantities within the model that will reappear frequently throughout the thesis.

2.1 Biological background

2.1.1 Coral-algal symbiosis

Stony reef-building corals, such as those found on the GBR, are comprised of two main species: the coral *polyps*, and a dinoflagellate unicellular algae of the family *Symbiodiniaceae* (commonly called *zooxanthellae*), which lives within and forms an important symbiosis with the coral polyps. This symbiosis is typically obligate, meaning that coral and algae cannot survive without one another. The organism formed as a result is called the *holobiont*, and it is due to this obligate symbiosis that the holobiont is able to grow, reproduce and form calcium carbonate structures so rapidly, outpacing the destruction of these structures due to erosion and boring organisms [71].

Energy is provided to the coral via photosynthetic products of its symbiont, *Symbiodiniaceae*, which fix and translocate carbon to their host, in the form of glycerol, glucose and amino acids: these are used by the coral to produce proteins, fats and carbohydrates [36, 62], which in turn provide energy to the host. The coral host receives up to 100% of its daily energy requirements from its symbiont [63], and can store excess energy in the form of lipids [40, 65]. In return, the symbiont receives protection, access to light (which is necessary to facilitate photosynthesis), and metabolic waste products from their host [32].

Coral reef structures are predominantly formed of white calcium carbonate, deposited by reef wildlife, or by the coral polyps themselves [60]; however, photosynthetic pigments in the algae cells give stony corals a deep brown colour [58]. *Symbiodiniaceae* cells are minuscule; one square centimetre of coral tissue tends to house approximately one million algal cells during ambient conditions [64].

2.1.2 Coral bleaching

Coral bleaching is a term given to the whitening of coral due to loss of resident *Symbiodiniaceae*. The physical colouration of a coral host is wholly derived from the photosynthetic pigments of resident *Symbiodiniaceae* which, when taken away, expose the coral's white calcium carbonate skeleton—hence, the term *bleaching*.

During bleaching events, the loss of *Symbiodiniaceae* leads to reduced energy production by the symbiont population. If the bleaching is severe enough, the symbiont population cannot produce enough energy to sustain the coral host, and the host must rely on stored lipids for survival [39]. Corals can depend on stored energy on the scale of weeks to months [27]; subsequently, if the host cannot replenish its symbiont population, it will die from energy starvation.

Recent evidence shows that symbiont 'switching' can occur in reef-building corals, potentially driven by consecutive bleaching events [16]. The term 'switching' refers to the coral host uptaking new *Symbiodiniaceae* from the environment, rather than the growth of an already existent population within the host, although this uptake is typically within juvenile corals only.

Coral bleaching is primarily induced by sustained heat stress due to high sea-surface temperatures (SST) coupled with high levels of ultraviolet (UV) radiation [44], but can be affected by other environmental factors including, but not limited to, salinity [47], pH [46], disease [17] and herbicides [53]. Coral bleaching is primarily due to the expulsion of *Symbiodiniaceae* from the coral host as a reaction to the stress caused by these environmental factors, rather than the death or loss of pigment from *Symbiodiniaceae* cells themselves [15].

It is not exactly clear what this 'reaction to stress' is. Bleaching is inexorably linked to the breakdown of the symbiosis between host and symbiont, but the mechanisms of how and why this happens on a cellular or molecular level are not entirely understood.

The currently accepted reason for this breakdown is that the stress caused by high temperature and UV radiation releases highly reactive chemicals in the symbiont cells. If the level of stress is high enough, these highly reactive chemicals start to move into the coral host, causing an immune-like response leading to the expulsion of the symbiont.

In more technical terms, high SST and UV damage the photosystem II[1] in *Symbiodiniaceae* cells, leading to the production of *reactive oxygen species* (ROS) such as superoxide, hydrogen peroxide, hydroxyl radicals and singlet oxygen [74]. ROS are typically regulated via antioxidant systems and other mechanisms within both the host and symbiont; however, under stress, these systems are either inhibited or cannot keep up with the level of ROS being produced [5].

Once spread to the coral host, ROS may trigger a response analogous to an immune response in the coral host, resulting in the expulsion of the symbiont from the host [33]. Baird *et al.* [5] alternatively suggests that when under stress, ROS production impairs the coral host's ability to distinguish between healthy and underperforming symbionts, leading to the expulsion of healthy symbionts to the detriment of its own health. This hypothesis is supported by Davy *et al.* [28].

Bleaching severity

Some bleaching events are more severe than others. Mildly bleached corals may recover on the scale of weeks to months [31], but severely bleached corals will starve and die if they are not able to recover enough symbiont density to sufficiently satisfy their energy demand [3].

Bleaching events can also be classified as either *local* (small-scale: a single colony or a single reef) or *mass* (large-scale: reef systems). Whereas small scale bleaching events can be influenced by a plethora of factors (SST, UV radiation, salinity, pH, disease, etc.) and are much more common, mass bleaching events are purely caused by high SST and UV radiation, and are far rarer [58]. Subsequently, the prediction of mass bleaching events is heavily reliant on SST/UV predictions.

[1] Photosystems I and II are protein structures which facilitate photosynthesis in various plants and algae.

Coral bleaching data

Due to the high level of interest in coral bleaching, there is an obvious need for data to monitor reefs, perform experiments, and conduct research. On reefs such as the GBR, data collection is a difficult problem, simply due to the immense area spanned by the reef. Data collection strategies can range from community monitoring, which is cost effective and covers large areas but is less precise, to research monitoring, which provides more reliable, accurate and precise data, but is expensive and only accounts for small localised areas [43].

Another factor in coral research is the type of data to collect, and the method of collection. Aerial and satellite surveys tend to be the most effective method for collecting data on large-scale environmental stress (i.e., the amount of bleaching in an area) [41]. There are various methods for measuring the degree of bleaching in coral, including the use of a colour reference card [69], and image processing techniques [19].

Non-invasive methods of collection, such as those mentioned above, are preferable over methods which disrupt coral habitats, however, methods of the latter variety are required to directly measure quantities such as symbiont density, relative symbiont abundances and carbon content which are more useful in understanding the mechanisms behind coral bleaching.

2.1.3 Genera of *Symbiodiniaceae*

Until recently, the symbiotic micro-algae that inhabit reef corals was known as the genus *Symbiodinium*, split into several distinct *clades*[2]. This classification has recently been revised, such that these different clades are now considered different *genera*[3] of the family *Symbiodiniaceae* [55].

Previously, different clades were denoted by arbitrary capital letters (A, B, C, etc.). The new naming convention introduced in LaJeunesse *et al.* [55] gives unique names to each genus of *Symbiodiniaceae*.

The most common genera of *Symbiodiniaceae* on the GBR are *S. Cladocopium* and *S. Durusdinium* (formerly clades C and D, respectively) [59]. A 2007 study showed that many corals on the GBR that contain *S. Cladocopium* also contain *S. Durusdinium* [59]. In general, coral hosts are known to form symbioses with multiple genera of symbiont at the same time, and over time [7, 56].

[2]The term *clade* refers to an evolutionary lineage of a species, distinguished by morphological, anatomical, and ecological features.

[3]Plural of genus, a taxonomic rank between family and species.

Different combinations of coral host and symbiont have different environmental tolerances and preferences. In particular, corals harbouring *S. Durusdinium* have been shown to be more thermally tolerant than corals harbouring *S. Cladocopium* [68].

S. Cladocopium tends to be the more favourable partner under non-stressful temperatures, due to its higher photochemical efficiency [22, 25] and enhanced ability to acquire nitrogen at these temperatures [9].

The interplay between the host and different genera of *Symbiodiniaceae* is crucial to understanding coral-algal symbiosis, due to the potential differences in growth, bleaching tolerance and energy production associated with the combination of the host and different symbiont genera.

2.1.4 The adaptive bleaching hypothesis

The adaptive bleaching hypothesis (ABH) is a theory conceived in 1993 [18]. The hypothesis was concisely re-stated in a 2004 paper [35] as follows.

> "[The ABH is] the proposal that the loss of photosymbionts has the potential to allow some representatives of the host species to re-establish a symbiosis with a different dominant alga, resulting in a new holobiont (also referred to as ecospecies or host-symbiont unit) that is better suited to the altered environmental circumstances."

In this statement, the 'loss of photosymbiont' implicitly refers to the expulsion of *Symbiodiniaceae* due to coral bleaching, thus the hypothesis insinuates that bleaching allows the coral host to be recolonised by a better adapted symbiont (i.e., a genus of *Symbiodiniaceae* better adapted to the current environment).

The ABH received both support [8] and criticism [37, 45] throughout the 2000s, but is no longer a prominent focus in coral literature due to advances made in cell biology which somewhat antiquate the ABH. The mysterious mechanism of coral bleaching is now thought to be for the purposes of defence (from damaging ROS) rather than adaptation.

Despite the ABH losing its lustre in the world of coral research, it is still notable due to the connection with the model on which this thesis is founded. The original model, which aimed to use a mathematical model to provide evidence for or against the ABH, is covered in more detail in Chapter 3.

2.1.5 Research questions

The ABH was the original motivation for the model in Chapter 3, but is no longer a research question we would like to answer; however, there are many other avenues to pursue in the field of coral modelling.

A particular avenue concerns the different genera of *Symbiodiniaceae*. It is apparent that *S. Durusdinium* is a more thermally tolerant genus, but whether it is always better for coral to host *S. Durusdinium* only is unclear. Perhaps the ability to host multiple genera of *Symbiodiniaceae* is an advantage, or perhaps it hinders the host, and the ability to host only a single genus is advantageous. This is a question we aim to answer with the help of a model.

Although the mechanisms of coral bleaching are not fully understood, a model may be able to illuminate these mechanisms without the need to collect data such as symbiont densities over time, which can be onerous.

2.2 Literature review

Before delving into our own model, we review previous models for coral bleaching and coral-algal symbiosis. This review is not exhaustive, but aims to provide an overview of the kind of models that exist and the features of coral bleaching that have been modelled, and set the scene for a new model.

2.2.1 Models for coral bleaching

These models focus on modelling the phenomenon of coral bleaching on a large scale, typically using collected data to predict future coral bleaching events or the long term effects of coral bleaching.

Berkelmans *et al.* [14] employs a statistical model, using SST data from mass-bleaching events in 1998 and 2002 to predict bleaching on the GBR. The approach also creates models which included a variable for cloud cover, area, perimeter, reef type and water depth. The model is able to correctly predict whether bleaching occurred or not on 90% of reefs. In particular, the maximum SST occurring over a 3-day period, denoted *max3d*, was found to be a good predictor of future bleaching occurrence.

Wooldridge and Done [77] applies a Bayesian approach to predict coral mortality on the GBR, also using SST data from a 2002 mass-bleaching event. The approach uses the *max3d* proxy from Berkelmans *et al.* [14] to great success, illustrating that proxy indicators can be very useful in prediction models for coral bleaching (as opposed to using many predictors from a dataset).

Van Hooidonk *et al.* [72] uses an ensemble of climate models, under four Representative Concentration Pathways[4] (RCPs), to predict the percentage of bleached coral reefs in the future. Even under the most favourable RCP, the model predicts that globally, most coral reefs will experience annual bleaching events by 2100.

Logan *et al.* [57] applies a similar approach, using SST data to predict bleaching prevalence, with the addition of consideration for adaptation from the coral host. The study focuses on bleaching outcomes under the 'moderately high' and 'highest' RCPs, predicting the percentage of reefs that will experience high-frequency bleaching in each year until 2100. The adaptive models lead to less bleached reefs in most cases, the degree of which depends on the assumed RCP.

2.2.2 Models for coral-algal symbiosis

These models focus on coral bleaching at a much smaller scale, that is, they explicitly model the symbiotic relationship between coral and algae. We are more interested in these types of models, as they look at the small-scale dynamics and mechanisms of coral bleaching.

Ware *et al.* [73] is a notable model, due to its connection with the Adaptive Bleaching Hypothesis. The model uses a differential equations (DEs) approach to model the prevalence of various clades of zooxanthellae within a coral colony under various degrees of temperature stress. The model attempts to incorporate adaptive behaviour, assuming that stress-sensitive combinations of coral and algae have advantages in non-stressful situations, and stress-tolerant combinations are advantageous in stressful situations. A major drawback to the model is that it does not account for coral mortality, which occurs alongside bleaching. Although more sophisticated models have come along since, it is one of the earliest mechanistic models for coral-algal symbiosis, and contains features which will be revisited later in this thesis.

Another notable set of models is the so-called class of Dynamic Energy Budget (DEB) models: Muller *et al.* [61] is the first example of DEB models being applied to coral-algal symbiosis: the method consists of a series of DEs which track various processes in the symbiosis, such as carbon and nitrogen transfer, and light irradiance. Eynaud *et al.* [34] studies this model in further detail, looking at how increased radiation is detrimental to the symbiosis.

[4] Representative Concentration Pathways are trajectories of greenhouse gas emissions until the year 2100, under various degrees of human involvement in reducing emissions. RCP 2.6 corresponds to emissions being drastically reduced by 2020, whereas RCP 8.5 corresponds to emissions continuing to grow through to 2100.

Cunning *et al.* [24] is a recent improvement of the model from Muller *et al.* [61]; it models oxidative stress explicitly, and is less complex than its predecessor.

Baird *et al.* [6] is one of the most recent mechanistic models for coral-algal symbiosis, again using a DE approach to model various processes encompassed in coral-algal symbiosis. The model accounts for nutrient sharing between host and symbiont, ROS production and photoadaptation (ability of symbiont to adapt to ROS production). In addition, the authors used the model, along with environmental data from the GBR to predict bleaching on the GBR, to varying degrees of success. The ability of this model to operate on several spatial scales is unique in the scope of models considered here.

Drawbacks

We are primarily interested in models for the mechanism of coral-algal symbiosis, rather than models created to predict coral bleaching, so we will discuss the drawbacks of the mechanistic models only.

One drawback is model complexity: the model from Baird *et al.* [6] in particular is quite complex, and may be less accessible to biologists as a result.

More importantly, in each of the mechanistic models for coral-algal symbiosis we have seen here, a common element is missing: *randomness*. All of the aforementioned mechanistic models use a deterministic DE approach, and as such do not capture the varying environmental and ecological factors which influence coral bleaching. To better understand the mechanisms of coral bleaching, a *stochastic* model—which can account for these varying factors—may be more appropriate.

2.3 The Abate-Whitt framework

Laplace transforms (LTs) [76] are a widely-used integral transform with applications in signal processing, electronic engineering and, of course, mathematics. The unilateral LT of a function $f(t)$ is defined for $t \in [0, \infty)$ and $s \in \mathbb{C}$ as

$$F(s) := \mathcal{L}\{f(t)\}(s) = \int_0^\infty e^{-st} f(t)\, dt.$$

We consider functions of the form $f : \mathcal{U} \to \mathbb{R}$, where $\mathcal{U} \subset \mathbb{R}$ is called the *time domain*. The LT is of the form $F : \mathcal{V} \to \mathbb{C}$, where $\mathcal{V} \subset \mathbb{C}$ is called the *frequency domain*.

A common problem encountered when dealing with LTs is the inversion of the transform. That is, one has obtained the transform $F(s)$, and would like to recover the function $f(t) = \mathcal{L}^{-1}\{F(s)\}(t)$.

For $u \in \mathcal{U}$, the inverse transform is defined by the Bromwich integral formula:

$$f(u) = \frac{1}{2\pi i} \lim_{L\to\infty} \int_{\gamma-iL}^{\gamma+iL} e^{su} F(s)\, ds, \tag{2.1}$$

where γ can be *any* real constant greater than the real part of all singularities of $e^{su}F(s)$ in $\mathbb{C}$. If the LT is too complicated, this formula cannot be used directly.

In these cases, there are methods for inverting LTs *numerically*, several of which come under the so-called *Abate-Whitt framework*, a general mathematical formulation of the inverse LT. The rest of this section explores the use of algorithms within this framework to approximate the inverse in Eq. (2.1).

2.3.1 The framework

The Abate-Whitt framework was introduced in 2006 [2], and provides a generic formula for inversion with the advantage that previous methods of inversion, such as the Euler method [1], can be written in terms of this framework. The framework can be derived in two ways: the first uses the Dirac delta functional (or generalised function), and the latter is derived from the Bromwich integral formula.

Consider a function $f(t)$ with LT $F(s)$. Then for $u \in \mathcal{U}$,

$$f(u) = \int_0^\infty \delta(t-u) f(t)\, dt, \quad u > 0 \tag{2.2}$$

where $\delta(x)$ is the *Dirac delta functional* [75]. This functional can be approximated by a weighted sum of exponential terms [67]:

$$\delta(t-u) \approx \delta_n(t-u) = \frac{1}{u} \sum_{k=1}^{n} \eta_k e^{-\beta_k t/u}, \tag{2.3}$$

where β_k are called the *nodes*, and η_k are called the *weights*. The number of nodes and weights, n, is called the *order* of the approximation.

Substituting Eq. (2.3) into Eq. (2.2), we obtain

$$\begin{aligned} f(u) &\approx \int_0^\infty \frac{1}{u} \sum_{k=1}^{n} \eta_k e^{-\frac{\beta_k}{u}t} f(t)\, dt \\ &= \frac{1}{u} \sum_{k=1}^{n} \eta_k \int_0^\infty e^{-\frac{\beta_k}{u}t} f(t)\, dt \\ &= \frac{1}{u} \sum_{k=1}^{n} \eta_k F\left(\frac{\beta_k}{u}\right). \end{aligned} \tag{2.4}$$

Thus, the inverse function can be approximated by a weighted sum of transform values. Importantly, the nodes and weights do not depend on $F(s)$ or on the point $u \in \mathcal{U}$ at which we would like to evaluate f. Therefore, the problem of inverting $F(s)$ is reduced to approximating the Dirac delta functional as in Eq. (2.3), by choosing appropriate values for β_k, η_k and n.

The same result can be found by considering the Bromwich integral formula, as in Abate and Whitt [2]. Substituting $z = su$ in Eq. (2.1), we obtain

$$f(u) = \frac{1}{2\pi i u} \lim_{L\to\infty} \int_{\gamma-iL}^{\gamma+iL} e^z F\left(\frac{z}{u}\right) dz. \tag{2.5}$$

The exponential term can be approximated[5] by a sum of rational functions:

$$e^z \approx \sum_{k=1}^{n} \frac{\eta_k}{\beta_k - z},$$

for some complex numbers η_k and β_k. Substituting this approximation into Eq. (2.5), we obtain the n^{th} order approximation for $f(u)$:

$$\begin{aligned} f_n(u) &= \frac{1}{2\pi i u} \lim_{L\to\infty} \int_{\gamma-iL}^{\gamma+iL} \sum_{k=1}^{n} \frac{\eta_k}{\beta_k - z} F\left(\frac{z}{u}\right) dz \\ &= \frac{1}{u} \sum_{k=1}^{n} \eta_k \left(\frac{1}{2\pi i} \lim_{L\to\infty} \int_{\gamma-iL}^{\gamma+iL} \frac{F(z/u)}{\beta_k - z}\, dz \right) \\ &= \frac{1}{u} \sum_{k=1}^{n} \eta_k \left(\frac{1}{2\pi i} \lim_{L\to\infty} \int_{\gamma-iL}^{\gamma+iL} \frac{F(z/u)}{(\beta_k/u) - (z/u)} \times \frac{1}{u}\, dz \right) \\ &= \frac{1}{u} \sum_{k=1}^{n} \eta_k \left(\frac{1}{2\pi i} \lim_{L\to\infty} \int_{(\gamma-iL)/u}^{(\gamma+iL)/u} \frac{F(s)}{a_k - s}\, ds \right), \end{aligned} \tag{2.6}$$

making the substitutions $s = z/u$, $a_k = \beta_k/u$, using the fact that $ds = (1/u)\, dz$. Using the Cauchy integral formula in Eq. (2.6), we obtain the expression

$$f_n(u) = \frac{1}{u} \sum_{k=1}^{n} \eta_k F\left(\frac{\beta_k}{u}\right),$$

which is exactly Eq. (2.4).

[5] In earlier methods for Laplace transform inversion, a common method of approximating the exponential function in this manner was the *Padé approximation*. Rodrigues [67] uses a Padé approximant to the Dirac delta, and shows that this approximation contains oscillations to the right of the centre of the Dirac delta. This results in oscillations to the right of discontinuities in inverted functions, similar to the oscillations in the Euler approximation in Figure 2.1.

2.3.2 The Euler method

The Euler algorithm is a popular method for inverting LTs. Originally introduced in a 1995 paper [1], the method is a modified version of a Fourier series algorithm for LT inversion [23], using Euler summation to speed up convergence.

Abate and Whitt [2] shows that the method fits into the Abate-Whitt framework, using the following parameter values:

$$\begin{aligned}\beta_k &= \frac{M\ln(10)}{3} + \pi i k,\\ \eta_k &= 10^{M/3}(-1)^k \xi_k,\end{aligned}$$

where $M = (n-1)/2$ (n must be odd), and

$$\xi_k = \begin{cases} \frac{1}{2} & k = 1,\\ 1 & 2 \le k \le M+1,\\ \xi_{k-1} - 2^{-M}\binom{M}{2M-k} & M+1 < k \le 2M,\\ 2^{-M} & k = 2M+1.\end{cases}$$

The method was originally derived from the Bromwich integral, which is why it fits into the Abate-Whitt framework. In terms of Eq. (2.1), the Euler method is equivalent to integrating over the vertical line $\gamma = M\ln(10)/3u$ in $\mathbb{C}$.

The Euler method is a common method for numerical LT inversion, but does not perform particularly well when applied to LTs of functions with discontinuities, such as the staircase function in Figure 2.1. We explore this phenomenon in more detail in Chapter 4, as discontinuous functions become particularly relevant to our investigation.

2.3.3 The Gaver-Stehfest method

The Gaver-Stehfest method was the other algorithm to appear in the original paper on the Abate-Whitt framework [2]. The algorithm, similar to the Euler algorithm, uses a form of convergence acceleration. The algorithm uses the following formulae to evaluate the weights and nodes, for $1 \le k \le n$:

$$\begin{aligned}\beta_k &= k\log(2),\\ \eta_k &= (-1)^{n/2+k}\log(2)\sum_{j=\lfloor (k+1)/2\rfloor}^{\min(k,n/2)} \frac{j^{(n/2)+1}}{(n/2)!}\binom{n/2}{j}\binom{2j}{j}\binom{j}{k-j}.\end{aligned}$$

The Gaver-Stehfest method is alluring since the nodes and weights do not involve complex numbers; however, modern methods such as the CME method (below) far surpass the Gaver-Stehfest method in accuracy. Figure 2.1 illustrates how the method struggles to handle functions with discontinuities.

2.3.4 The CME method

The Concentrated Matrix Exponential (CME) method is a recently developed method for LT inversion using Matrix Exponential (ME) distributions [48, 49, 51]. Incidentally, the probability density function (PDF) of a ME distribution is in the form displayed in Eq. (2.3), and thus an ME distribution can be used to approximate the Dirac delta functional. The real power of the method is the use of numerical optimisation to find a ME distribution with the lowest possible variance [50], for any given order. Here, *order* again refers to the number of nodes and weights in Eq. (2.4), the Abate-Whitt formula.

When applied to discontinuous functions such as the staircase function, the CME method outperforms both the Euler and Gaver-Stehfest method, as shown in Figure 2.1. A particularly important feature of the method is that it does not *overshoot*. In the context of functions $f(t)$ with some discontinuity at a point a, this means that the value of the CME approximation around a will be between $f(a-\varepsilon)$ and $f(a+\varepsilon)$ for some small $\varepsilon > 0$. In contrast, the Euler method can lead to overshooting, illustrated in Figure 2.1, where the Euler approximation (orange line) lies above the true function (blue line) at $t = 3.5$.

The non-overshooting property comes from the fact that the ME distribution is non-negative (it is a probability distribution), and thus the Dirac delta approximation is non-negative. If we consider the expression for the n^{th} order approximation $f_n(u)$ of the inverse transform $f(u)$,

$$f_n(u) = \int_0^\infty f(t)\delta_n(t-u)\,dt,$$

where $\delta_n(t-u)$ is the ME approximation, then we see that $f_n(u)$ is a weighted integral of $f(t)$ around u. If the ME approximation is perfect, then u will have all the weight, and we recover $f(u)$ exactly. In practice, the approximation gives weight to values in some interval around u, and thus we recover a value in some interval around $f(u)$. The size of this interval will depend on how good the ME approximation is (therein lies the reason for *concentrated* ME distributions). Even if all the 'weight' is centred on a value v far away from u, this will not lead to overshooting, since $f_n(u)$ will not exceed $f(v)$. This holds for any order n.

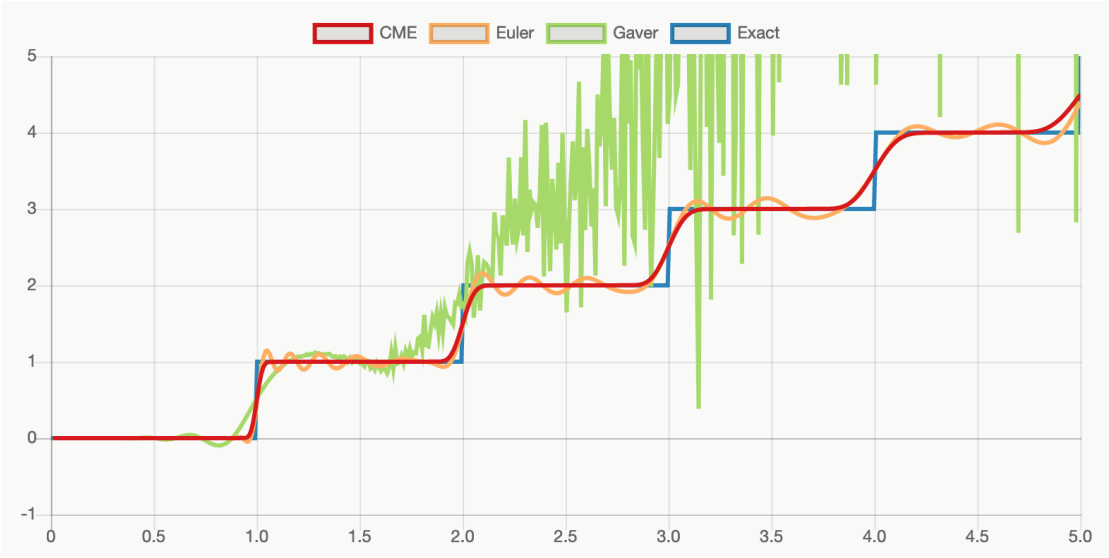


Figure 2.1: Comparison of Euler, Gaver-Stehfest and CME methods for inverting the Laplace transform of the staircase function. Plot created using the online tool available at `http://inverselaplace.org`.

Another important property of the CME method is that for an order N approximation, the squared coefficient of variation (SCV, comparable to variance) of the ME approximation is approximately $2/N^2$ [50]. Thus, the ME approximation always improves (the SCV decreases) as the order increases.

2.4 Stochastic fluid flow models

Stochastic fluid flow models, also known as Markov-modulated fluid models or fluid queues, are versatile models that consider a fluid buffer (e.g., a rainwater tank), and the level of fluid in this buffer (e.g., the level of rainwater). In a Markov-modulated fluid model, the rate at which fluid enters and leaves this buffer is determined by an irreducible Markov chain. We use Markov-modulated fluid models to model the mechanics of coral-algal symbiosis in the sequel, and so we introduce them in detail here.

2.4.1 Unbounded fluid flows

An unbounded Markov-modulated fluid flow model is a two-dimensional stochastic process $\{\mathcal{F}(t)\}_{t\geq 0} = \{(M(t), \varphi(t)) \in \mathbb{R} \times \mathcal{S} : t \geq 0\}$ where $M(t)$ denotes the *level* and $\varphi(t)$ denotes the *phase*. The phase process is an irreducible continuous-time Markov chain (CTMC) with infinitesimal generator $\mathbf{T}$ and state space $\mathcal{S}$.

To each phase $i \in \mathcal{S}$ is attributed a fluid input rate $c_i \in \mathbb{R}$, and thus, the state space can be partitioned as $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_- \cup \mathcal{S}_0$, such that:

- $\mathcal{S}_+ = \{i \in \mathcal{S} : c_i > 0\}$;
- $\mathcal{S}_- = \{i \in \mathcal{S} : c_i < 0\}$; and
- $\mathcal{S}_0 = \{i \in \mathcal{S} : c_i = 0\}$.

In addition, let $s_+ = |\mathcal{S}_+|$, $s_- = |\mathcal{S}_-|$ and $s_0 = |\mathcal{S}_0|$.

Define the matrix $\mathbf{C}$ as the diagonal matrix of non-zero fluid input rates, such that $\mathbf{C}_{i,i} = |c_i|$ for $i \in \mathcal{S}_+ \cup \mathcal{S}_-$ (hence $\mathbf{C}$ has positive diagonal entries).

We partition $\mathbf{T}$ and $\mathbf{C}$ according to $\mathcal{S}_+$, $\mathcal{S}_-$ and $\mathcal{S}_0$, as appropriate:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{++} & \mathbf{T}_{+-} & \mathbf{T}_{+0} \\ \mathbf{T}_{-+} & \mathbf{T}_{--} & \mathbf{T}_{-0} \\ \mathbf{T}_{0+} & \mathbf{T}_{0-} & \mathbf{T}_{00} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_+ & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_- \end{bmatrix}.$$

The behaviour of the level process, $M(t)$, is entirely determined by the phase process, $\varphi(t)$. When the phase process is in state $i \in \mathcal{S}$, the level changes at a constant rate c_i. The level increases when $i \in \mathcal{S}_+$, decreases when $i \in \mathcal{S}_-$, and is constant when $i \in \mathcal{S}_0$. Mathematically, the level process can be defined as

$$M(t) = M(0) + \int_0^t c_{\varphi(u)}\, du.$$

Figure 2.2 illustrates how the level and phase processes evolve over time.

A common quantity of interest in fluid models is the matrix $\mathbf{\Psi}$, as well as its counterpart $\mathbf{\Xi}$, which give the probability of returning to an initial level z: $\mathbf{\Psi}$ is used for paths which begin in an upwards phase, and $\mathbf{\Xi}$ is used for paths which begin in a downwards phase.

Define $\theta(x) = \inf\{t > 0 : M(t) = x\}$. Then, for $i \in \mathcal{S}_+$ and $j \in \mathcal{S}_-$, we define

$$\mathbf{\Psi}_{i,j} = \mathbb{P}[\theta(z) < \infty, \varphi(\theta(z)) = j \mid M(0) = z, \varphi(0) = i], \tag{2.7}$$

$$\mathbf{\Xi}_{j,i} = \mathbb{P}[\theta(z) < \infty, \varphi(\theta(z)) = i \mid M(0) = z, \varphi(0) = j]. \tag{2.8}$$

Note that the quantities $\mathbf{\Psi}$ and $\mathbf{\Xi}$ are independent of the initial level z.

Paths in the unbounded and bounded models

Throughout this thesis, we frequently use *path-based* arguments when working with fluid models. A *path* is any single realisation of the fluid model (i.e., the

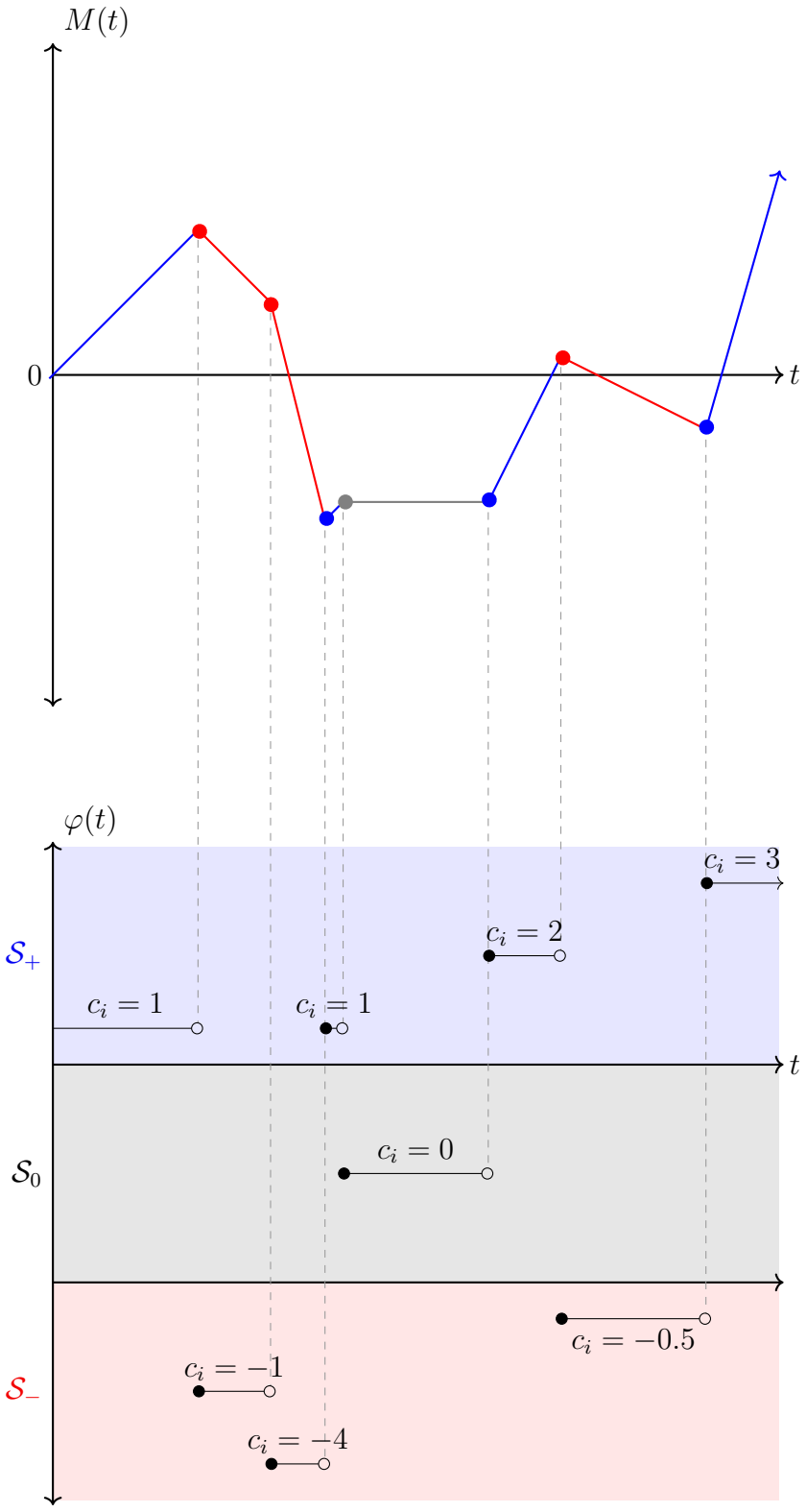


Figure 2.2: An example realisation of the unbounded fluid flow process $\{\mathcal{F}(t)\}_{t\geq 0}$. The phase process $\varphi(t)$ evolves according to a continuous-time Markov chain, and the level process $M(t)$ increases/decreases at rate c_i when the current phase is i. Here, the initial level $M(0) = 0$, but this can be any value.

level *and* phase processes). Path-based arguments become particularly important when we begin to derive the so-called *time to mortality* in later chapters.

In a fluid model (bounded or unbounded), we can classify any path into one of the following four categories:

- Up-up (++): the fluid model begins in phase $i \in \mathcal{S}_+$ and ends in $j \in \mathcal{S}_+$. This type of path is illustrated in Figure 2.3 (d) and (f).
- Up-down (+−): the fluid model begins in phase $i \in \mathcal{S}_+$ and ends in $j \in \mathcal{S}_-$. This type of path is illustrated in Figure 2.3 (a), (c) and (e).
- Down-up (−+): the fluid model begins in phase $i \in \mathcal{S}_-$ and ends in $j \in \mathcal{S}_+$. This type of path is illustrated in Figure 2.3 (b), (d) and (f).
- Down-down (−−): the fluid model begins in phase $i \in \mathcal{S}_-$ and ends in $j \in \mathcal{S}_-$. This type of path is illustrated in Figure 2.3 (c) and (e).

To concatenate paths, we use the fact that paths are conditionally independent of each other, given the phase. Over the page, the insert Concatenating paths uses an example to explain this concept further.

Time-dependent quantities in unbounded fluid flows

Bean *et al.* [11, 13] derives expressions for important quantities which account for time by using Laplace-Stieltjes transforms (LSTs). Throughout this section, it is implicitly assumed that $\Re[s] \geq 0$. A summary of the quantities derived in this section is provided in Figure 2.3, which illustrates sample paths for each quantity.

The first important quantities are $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$, which denote the LSTs of the time taken to return to an initial level z ($\widehat{\mathbf{\Psi}}(s)$ for paths which begin in upward phases, $\widehat{\mathbf{\Xi}}(s)$ for paths which begin in downward phases). For $i \in \mathcal{S}_+$, $j \in \mathcal{S}_-$ and $z \in \mathbb{R}$, we define

$$\left[\widehat{\mathbf{\Psi}}(s)\right]_{i,j} = \mathbb{E}\left[e^{-s\theta(z)}\mathbb{I}(\theta(z) < \infty, \varphi(\theta(z)) = j) \;\middle|\; M(0) = z, \varphi(0) = i\right], \tag{2.9}$$

$$\left[\widehat{\mathbf{\Xi}}(s)\right]_{j,i} = \mathbb{E}\left[e^{-s\theta(z)}\mathbb{I}(\theta(z) < \infty, \varphi(\theta(z)) = i) \;\middle|\; M(0) = z, \varphi(0) = j\right], \tag{2.10}$$

again, independent of the initial level z. Bean *et al.* [11] shows that $\widehat{\mathbf{\Psi}}(s)$ satisfies the algebraic Riccati equation

$$\mathbf{Q}_{+-}(s) + \widehat{\mathbf{\Psi}}(s)\mathbf{Q}_{-+}(s)\widehat{\mathbf{\Psi}}(s) + \mathbf{Q}_{++}(s)\widehat{\mathbf{\Psi}}(s) + \widehat{\mathbf{\Psi}}(s)\mathbf{Q}_{--}(s) = \mathbf{0}, \tag{2.11}$$

Concatenating paths using conditional independence

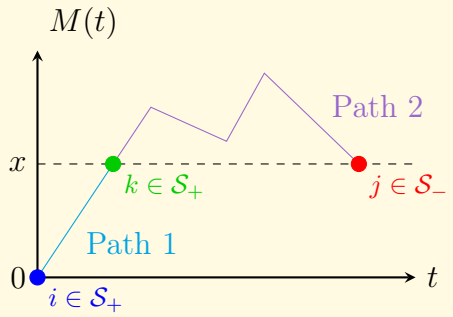
When dealing with fluid models, we often wish to construct complicated paths from simple paths. The four most simple paths are:

- Up-up ($++$, looks like $\nearrow$),
- Up-down ($+-$, looks like $\wedge$),
- Down-up ($-+$, looks like $\vee$), and
- Down-down ($--$, looks like $\searrow$).

Simple paths have a starting level and a destination level. The path is completed when the process reaches the destination level for the first time. The phase can change any number of times before the destination is reached.

As an example, consider two simple paths: an up-up path starting at level 0 with destination level x (Path 1), and an up-down path starting at level x, also with destination level x (Path 2). We would like to concatenate Path 1 and Path 2 to obtain a new path, Path 3.

Suppose Path 3 begins in phase $i \in \mathcal{S}_+$, and ends in $j \in \mathcal{S}_-$. During the journey to level x (Path 1), the phase process can make any number of transitions, and can end in any phase $k \in \mathcal{S}_+$. Once the process reaches level x in this phase, it can make any number of phase transitions, before returning to level x (Path 2) in phase $j \in \mathcal{S}_-$. The full journey might look like this:



full journey = Path 3

The important thing to note is that Path 1 and Path 2 are independent if the phase $k \in \mathcal{S}_+$ is pre-specified. We say that Path 1 and Path 2 are *conditionally independent* given the phase $k \in \mathcal{S}_+$.

We can account for *all* possible values of $k \in \mathcal{S}_+$ by using the law of total probability. This concept is revisited in Section 3.3.2, when we have introduced our problem in more detail.

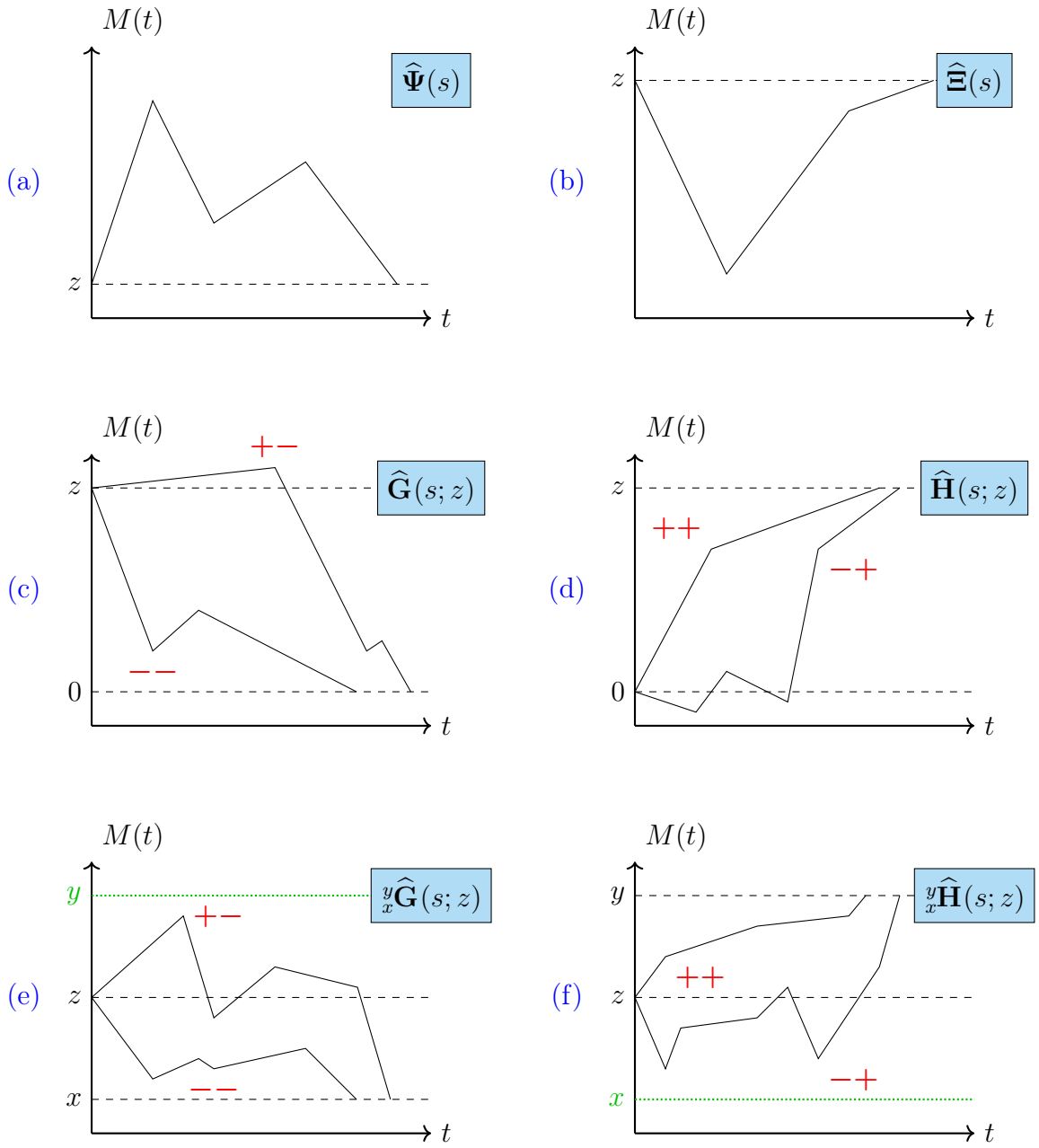


Figure 2.3: Example realisations of different paths in the unbounded fluid flow process $\{\mathcal{F}(t)\}_{t\geq 0}$, where z is the initial fluid level (except in $\widehat{\mathbf{H}}(s;z)$, where it is the destination level). Green dotted lines in (e) and (f) indicate taboo levels. Corresponding LSTs are displayed in the top right, which are the LSTs of the time taken to reach the destination level—and, in the case of $\overset{y}{x}\widehat{\mathbf{G}}(s;z)$ and $\overset{y}{x}\widehat{\mathbf{H}}(s;z)$, before reaching the taboo level.

where

$$\begin{aligned}
\mathbf{Q}_{++}(s) &= \mathbf{C}_{+}^{-1}\left[(\mathbf{T}_{++} - s\mathbf{I}) - \mathbf{T}_{+0}\left(\mathbf{T}_{00} - s\mathbf{I}\right)^{-1}\mathbf{T}_{0+}\right], \\
\mathbf{Q}_{--}(s) &= \mathbf{C}_{-}^{-1}\left[(\mathbf{T}_{--} - s\mathbf{I}) - \mathbf{T}_{-0}\left(\mathbf{T}_{00} - s\mathbf{I}\right)^{-1}\mathbf{T}_{0-}\right], \\
\mathbf{Q}_{+-}(s) &= \mathbf{C}_{+}^{-1}\left[\mathbf{T}_{+-} - \mathbf{T}_{+0}\left(\mathbf{T}_{00} - s\mathbf{I}\right)^{-1}\mathbf{T}_{0-}\right], \\
\mathbf{Q}_{-+}(s) &= \mathbf{C}_{-}^{-1}\left[\mathbf{T}_{-+} - \mathbf{T}_{-0}\left(\mathbf{T}_{00} - s\mathbf{I}\right)^{-1}\mathbf{T}_{0+}\right].
\end{aligned} \tag{2.12}$$

Further, $\widehat{\boldsymbol{\Xi}}(s)$ satisfies the symmetric equation:

$$\mathbf{Q}_{-+}(s) + \widehat{\boldsymbol{\Xi}}(s)\mathbf{Q}_{+-}(s)\widehat{\boldsymbol{\Xi}}(s) + \mathbf{Q}_{--}(s)\widehat{\boldsymbol{\Xi}}(s) + \widehat{\boldsymbol{\Xi}}(s)\mathbf{Q}_{++}(s) = \mathbf{0}. \tag{2.13}$$

Furthermore, if s is real, then $\widehat{\boldsymbol{\Psi}}(s)$ and $\widehat{\boldsymbol{\Xi}}(s)$ are the minimal non-negative solutions to their corresponding equations [11]. Solving the above Riccati equations is a non-trivial exercise, but there exist a plethora of algorithms for doing so [12].

Both $\widehat{\boldsymbol{\Psi}}(s)$ and $\widehat{\boldsymbol{\Xi}}(s)$ are very important, as they are used in calculating almost every time-dependent quantity that we shall use in this thesis. Example paths for $\widehat{\boldsymbol{\Psi}}(s)$ and $\widehat{\boldsymbol{\Xi}}(s)$ are illustrated in Figure 2.3 (a) and (b), respectively.

The next important quantities are $\widehat{\mathbf{G}}(s;z)$ and $\widehat{\mathbf{H}}(s;z)$, which are the LSTs of the time to respectively drain or fill the fluid buffer by an amount $z > 0$. For $i, j \in \mathcal{S}_+ \cup \mathcal{S}_-$, they are defined as

$$\left[\widehat{\mathbf{G}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\theta(0)}\mathbb{I}(\theta(0) < \infty, \varphi(\theta(0)) = j) \;\middle|\; M(0) = z, \varphi(0) = i\right], \tag{2.14}$$

$$\left[\widehat{\mathbf{H}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\theta(z)}\mathbb{I}(\theta(z) < \infty, \varphi(\theta(z)) = j) \;\middle|\; M(0) = 0, \varphi(0) = i\right]. \tag{2.15}$$

Note that both quantities depend on z; however, since $\{\mathcal{F}(t)\}_{t\geq 0}$ is level-independent, $\widehat{\mathbf{G}}(s;z)$ is also the LST of the time taken to drain the level from $z + c$ to c, for any $c \in \mathbb{R}$. Similarly, $\widehat{\mathbf{H}}(s;z)$ is the LST of the time taken to fill from c to $z + c$.

Note that when draining or filling the fluid buffer, the process may start in any phase, but must end in a downwards phase if draining and an upwards phase if filling. Thus, when applying the partition according to $\mathcal{S}_+ \cup \mathcal{S}_-$, we obtain

$$\widehat{\mathbf{G}}(s;z) = \begin{bmatrix} \mathbf{0} & \widehat{\mathbf{G}}_{+-}(s;z) \\ \mathbf{0} & \widehat{\mathbf{G}}_{--}(s;z) \end{bmatrix}, \quad \widehat{\mathbf{H}}(s;z) = \begin{bmatrix} \widehat{\mathbf{H}}_{++}(s;z) & \mathbf{0} \\ \widehat{\mathbf{H}}_{-+}(s;z) & \mathbf{0} \end{bmatrix}. \tag{2.16}$$

Theorem 1 from Bean *et al.* [13] shows that

$$\begin{aligned}\widehat{\mathbf{G}}_{+-}(s;z) &= \widehat{\mathbf{\Psi}}(s)\exp(\mathbf{B}(s)z),\\ \widehat{\mathbf{G}}_{--}(s;z) &= \exp(\mathbf{B}(s)z),\\ \widehat{\mathbf{H}}_{++}(s;z) &= \exp(\mathbf{A}(s)z),\\ \widehat{\mathbf{H}}_{-+}(s;z) &= \widehat{\mathbf{\Xi}}(s)\exp(\mathbf{A}(s)z),\end{aligned} \tag{2.17}$$

where $\exp(\mathbf{M})$ denotes the matrix exponential, and

$$\begin{aligned}\mathbf{A}(s) &= \mathbf{C}_+^{-1}\left[(\mathbf{T}_{++}-s\mathbf{I}) + \mathbf{T}_{+-}\widehat{\mathbf{\Xi}}(s) - \mathbf{T}_{+0}\left(\mathbf{T}_{00}-s\mathbf{I}\right)^{-1}\left[\mathbf{T}_{0+}+\mathbf{T}_{0-}\widehat{\mathbf{\Xi}}(s)\right]\right],\\ \mathbf{B}(s) &= \mathbf{C}_-^{-1}\left[(\mathbf{T}_{--}-s\mathbf{I}) + \mathbf{T}_{-+}\widehat{\mathbf{\Psi}}(s) - \mathbf{T}_{-0}\left(\mathbf{T}_{00}-s\mathbf{I}\right)^{-1}\left[\mathbf{T}_{0-}+\mathbf{T}_{0+}\widehat{\mathbf{\Psi}}(s)\right]\right]\end{aligned}$$

are the infinitesimal generators of the filling/draining processes, respectively.

Example paths for $\widehat{\mathbf{G}}(s;z)$ and $\widehat{\mathbf{H}}(s;z)$ are illustrated in Figure 2.3 (c) and (d), respectively.

Filling/draining times with a taboo level

The final important quantities in the unbounded model are the matrices ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$, which correspond to the LSTs of the time taken to respectively drain or fill the fluid buffer without hitting a taboo level. For phases $i,j \in \mathcal{S}_+ \cup \mathcal{S}_-$ and $x < z < y$, they are defined as

$$\left[{}_x^y\widehat{\mathbf{G}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\theta(x)}\mathbb{I}(\theta(x)<\theta(y),\varphi(\theta(x))=j) \;\middle|\; M(0)=z,\varphi(0)=i\right], \tag{2.18}$$

$$\left[{}_x^y\widehat{\mathbf{H}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\theta(y)}\mathbb{I}(\theta(y)<\theta(x),\varphi(\theta(y))=j) \;\middle|\; M(0)=z,\varphi(0)=i\right]. \tag{2.19}$$

We interpret ${}_x^y\widehat{\mathbf{G}}(s;z)$ as the LST of the time taken to drain the fluid buffer from z to x, in *taboo* of level y. Here, the word taboo means the process ignores paths that hit the taboo level. This is reflected by the condition $\theta(x) < \theta(y)$. The quantity ${}_x^y\widehat{\mathbf{H}}(s;z)$ is the LST of the time taken to fill the fluid buffer from z to y, in taboo of level x, with a similar interpretation.

Example paths for ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$ are illustrated in Figure 2.3 (e) and (f), respectively, with the taboo levels indicated by green dotted lines.

The matrix ${}_0^y\widehat{\mathbf{G}}(s;z)$ here is equivalent to $\widehat{\mathbf{G}}^{z,y}(s)$ from Bean *et al.* [13] (for more detail, see the insert A note on notation over the page). Also note that due to level independence, ${}_x^y\widehat{\mathbf{G}}(s;z) = {}_0^{y-x}\widehat{\mathbf{G}}(s;z-x)$ and ${}_x^y\widehat{\mathbf{H}}(s;z) = {}_0^{y-x}\widehat{\mathbf{H}}(s;z-x)$.

A note on notation

Here, we introduce a new notation for the LSTs of filling and draining times, a slight variation of the notation used in Bean *et al.* [13]. The new notation for ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$ can be interpreted as follows: the left hand scripts denote the interval over which the path operates; that is, the level must remain in the interval $[x,y]$. The argument z denotes the starting level. One of x, y and z is redundant due to level-independence, but all three are included for clarity.

In the case of ${}_x^y\widehat{\mathbf{G}}(s;z)$, the fluid buffer is being drained, so x is implicitly the destination level, and y the taboo level.

In the case of ${}_x^y\widehat{\mathbf{H}}(s;z)$, the fluid buffer is being filled, so y is implicitly the destination level and x is the taboo level.

Implicitly, we must have $x < z < y$. We may evaluate ${}_x^y\widehat{\mathbf{G}}(s;x)$, ${}_x^y\widehat{\mathbf{G}}(s;y)$, ${}_x^y\widehat{\mathbf{H}}(s;x)$ and ${}_x^y\widehat{\mathbf{H}}(s;y)$ by limiting arguments.

Note that unlike $\widehat{\mathbf{G}}(s;z)$, ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$, the z in $\widehat{\mathbf{H}}(s;z)$ does not represent the starting level (which is 0), but the destination level.

Theorem 2 from Bean *et al.* [13] shows that for $0 \le z \le y$,

$$\begin{bmatrix} {}_0^y\widehat{\mathbf{G}}(s;z) & {}_0^y\widehat{\mathbf{H}}(s;z) \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{G}}(s;z) & \widehat{\mathbf{H}}(s;y-z) \end{bmatrix} \begin{bmatrix} \mathbf{I} & \widehat{\mathbf{H}}(s;y) \\ \widehat{\mathbf{G}}(s;y) & \mathbf{I} \end{bmatrix}^{-1}. \tag{2.20}$$

This matrix inverse always exists for $\Re[s] > 0$, since under this condition the dominant eigenvalues of $\widehat{\mathbf{G}}(s;y)$ and $\widehat{\mathbf{H}}(s;y)$ are strictly less than 1, and exists at $s = 0$ if $M(t)$ is transient [13].

Eq. (2.20) is derived from the equations

$$\widehat{\mathbf{G}}(s;z) = {}_0^y\widehat{\mathbf{G}}(s;z) + {}_0^y\widehat{\mathbf{H}}(s;z)\widehat{\mathbf{G}}(s;y), \tag{2.21}$$

$$\widehat{\mathbf{H}}(s;y-z) = {}_0^y\widehat{\mathbf{H}}(s;z) + {}_0^y\widehat{\mathbf{G}}(s;z)\widehat{\mathbf{H}}(s;y). \tag{2.22}$$

which can be rearranged to obtain

$${}_0^y\widehat{\mathbf{G}}(s;z) = \widehat{\mathbf{G}}(s;z) - {}_0^y\widehat{\mathbf{H}}(s;z)\widehat{\mathbf{G}}(s;y), \tag{2.23}$$

$${}_0^y\widehat{\mathbf{H}}(s;z) = \widehat{\mathbf{H}}(s;y-z) - {}_0^y\widehat{\mathbf{G}}(s;z)\widehat{\mathbf{H}}(s;y). \tag{2.24}$$

Substituting Eq. (2.24) into Eq. (2.23) yields

$$\begin{aligned}{}_0^y\widehat{\mathbf{G}}(s;z) &= \widehat{\mathbf{G}}(s;z) - \left[\widehat{\mathbf{H}}(s;y-z) - {}_0^y\widehat{\mathbf{G}}(s;z)\widehat{\mathbf{H}}(s;y)\right]\widehat{\mathbf{G}}(s;y) \\ &= \widehat{\mathbf{G}}(s;z) - \widehat{\mathbf{H}}(s;y-z)\widehat{\mathbf{G}}(s;y) - {}_0^y\widehat{\mathbf{G}}(s;z)\widehat{\mathbf{H}}(s;y)\widehat{\mathbf{G}}(s;y),\end{aligned}$$

which implies

$${}_0^y\widehat{\mathbf{G}}(s;z)\left[\mathbf{I} - \widehat{\mathbf{H}}(s;y)\widehat{\mathbf{G}}(s;y)\right] = \widehat{\mathbf{G}}(s;z) - \widehat{\mathbf{H}}(s;y-z)\widehat{\mathbf{G}}(s;y).$$

Thus,

$${}_0^y\widehat{\mathbf{G}}(s;z) = \left[\widehat{\mathbf{G}}(s;z) - \widehat{\mathbf{H}}(s;y-z)\widehat{\mathbf{G}}(s;y)\right]\left(\mathbf{I} - \widehat{\mathbf{H}}(s;y)\widehat{\mathbf{G}}(s;y)\right)^{-1}.$$

Again, since the dominant eigenvalues of $\widehat{\mathbf{G}}(s;y)$ and $\widehat{\mathbf{H}}(s;y)$ are less than 1 when $\Re[s] > 0$, this matrix inverse always exists for $\Re[s] > 0$, and for $s = 0$ if $M(t)$ is transient. A similar result can be obtained for ${}_x^y\widehat{\mathbf{H}}(s;z)$ by substituting Eq. (2.23) into Eq. (2.24). By taking the appropriate sub-matrices of $\widehat{\mathbf{G}}(s;x)$ and $\widehat{\mathbf{H}}(s;x)$, one obtains the following set of equations:

$$\begin{aligned}
{}_0^y\widehat{\mathbf{G}}_{+-}(s;z) &= \left[\widehat{\mathbf{G}}_{+-}(s;z) - \widehat{\mathbf{H}}_{++}(s;y-z)\widehat{\mathbf{G}}_{+-}(s;y)\right]\left(\mathbf{I} - \widehat{\mathbf{H}}_{-+}(s;y)\widehat{\mathbf{G}}_{+-}(s;y)\right)^{-1}, \\
{}_0^y\widehat{\mathbf{G}}_{--}(s;z) &= \left[\widehat{\mathbf{G}}_{--}(s;z) - \widehat{\mathbf{H}}_{-+}(s;y-z)\widehat{\mathbf{G}}_{+-}(s;y)\right]\left(\mathbf{I} - \widehat{\mathbf{H}}_{-+}(s;y)\widehat{\mathbf{G}}_{+-}(s;y)\right)^{-1}, \\
{}_0^y\widehat{\mathbf{H}}_{++}(s;z) &= \left[\widehat{\mathbf{H}}_{++}(s;y-z) - \widehat{\mathbf{G}}_{+-}(s;z)\widehat{\mathbf{H}}_{-+}(s;y)\right]\left(\mathbf{I} - \widehat{\mathbf{G}}_{+-}(s;y)\widehat{\mathbf{H}}_{-+}(s;y)\right)^{-1}, \\
{}_0^y\widehat{\mathbf{H}}_{-+}(s;z) &= \left[\widehat{\mathbf{H}}_{-+}(s;y-z) - \widehat{\mathbf{G}}_{--}(s;z)\widehat{\mathbf{H}}_{-+}(s;y)\right]\left(\mathbf{I} - \widehat{\mathbf{G}}_{+-}(s;y)\widehat{\mathbf{H}}_{-+}(s;y)\right)^{-1}.
\end{aligned} \tag{2.25}$$

This is true only because $\widehat{\mathbf{G}}_{++}(s;x) = \widehat{\mathbf{G}}_{-+}(s;x) = \mathbf{0}$ (must end in a negative phase if draining) and $\widehat{\mathbf{H}}_{+-}(s;x) = \widehat{\mathbf{H}}_{--}(s;x) = \mathbf{0}$ (must end in a positive phase if filling). Incidentally, it is also true that ${}_0^y\widehat{\mathbf{G}}_{++}(s;z) = {}_0^y\widehat{\mathbf{G}}_{-+}(s;z) = \mathbf{0}$ and ${}_0^y\widehat{\mathbf{H}}_{+-}(s;z) = {}_0^y\widehat{\mathbf{H}}_{--}(s;z) = \mathbf{0}$.

2.4.2 Bounded fluid flows

This section introduces the *bounded* stochastic fluid model, and explains Theorems 3 and 4 from Bean *et al.* [13], which give rise to important quantities which shall be used in later chapters. A bounded stochastic fluid model is very similar to the unbounded variant, except the level process is now restricted to a fixed interval $[0, b]$. The process has special behaviour at the lower and upper boundaries.

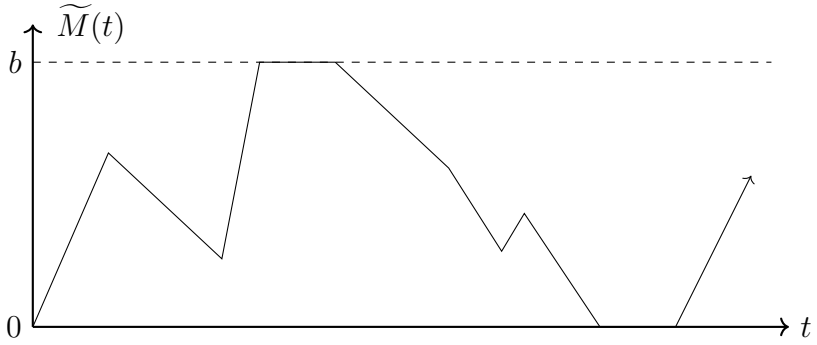


Figure 2.4: An example realisation of the bounded level process $\widetilde{M}(t)$.

The bounded process is defined to be the two-dimensional stochastic process $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0} = \{(\widetilde{M}(t), \widetilde{\varphi}(t)) \in [0,b] \times \mathcal{S} \cup \widehat{\mathcal{S}} \cup \check{\mathcal{S}} : t \geq 0\}$ where $\widetilde{M}(t)$ is the level and $\widetilde{\varphi}(t)$ the phase. Once again, $\mathbf{T}$ is the infinitesimal generator of the phase process $\widetilde{\varphi}(t)$ and $\mathbf{C}$ is a diagonal matrix containing the fluid input/output rates. $\widehat{\mathcal{S}}$ and $\check{\mathcal{S}}$ are the sets of phases at the upper and lower boundaries, respectively.

Within the interval $(0, b)$, the process evolves in the same manner as the unbounded model. To describe the behaviour of the process at the boundaries, we introduce the matrices $\widehat{\mathbf{P}}$, $\widehat{\mathbf{T}}$, $\check{\mathbf{P}}$ and $\check{\mathbf{T}}$.

$\widehat{\mathbf{P}}$ is the probability matrix governing transitions from $\mathcal{S}_+$ to $\widehat{\mathcal{S}} \cup \mathcal{S}_-$ when the level hits the upper boundary. Similarly, $\check{\mathbf{P}}$ is the probability matrix governing transitions from $\mathcal{S}_-$ to $\check{\mathcal{S}} \cup \mathcal{S}_+$ when the level hits the lower boundary.

$\widehat{\mathbf{T}}$ is the infinitesimal generator of the phase process at the upper boundary, and operates on the state space $\widehat{\mathcal{S}} \cup \mathcal{S}_-$. Similarly, $\check{\mathbf{T}}$ is the generator at the lower boundary, and operates on $\check{\mathcal{S}} \cup \mathcal{S}_+$. When the process enters $\mathcal{S}_-$ or $\mathcal{S}_+$, the process leaves the respective boundary and regular behaviour resumes.

The process behaviour at the boundaries is described explicitly as follows:

- **Upper boundary**: when the process hits level b in phase $i \in \mathcal{S}_+$, the phase process immediately transitions to a new phase $j \in \widehat{\mathcal{S}} \cup \mathcal{S}_-$ according to $\widehat{\mathbf{P}}_{i,j}$. If $j \in \mathcal{S}_-$, the process reflects off the boundary and behaviour resumes as normal. If $j \in \widehat{\mathcal{S}}$, the phase process evolves according to $\widehat{\mathbf{T}}$, until the phase enters $\mathcal{S}_-$. Until that occurs, the level remains at b.

- **Lower boundary**: when the process hits level 0 in phase $i \in \mathcal{S}_-$, the phase process immediately transitions to a new phase $j \in \check{\mathcal{S}} \cup \mathcal{S}_+$ according to $\check{\mathbf{P}}_{i,j}$. If $j \in \mathcal{S}_+$, the process reflects off the boundary and behaviour resumes as normal. If $j \in \check{\mathcal{S}}$, the phase process evolves according to $\check{\mathbf{T}}$, until the phase enters $\mathcal{S}_+$. Until that occurs, the level remains at 0.

An example realisation of $\widetilde{M}(t)$ is illustrated in Figure 2.4.

Time-dependent quantities in the bounded model

We would like to derive analogous quantities to $\widehat{\mathbf{\Psi}}(s)$, $\widehat{\mathbf{\Xi}}(s)$, ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$ in the bounded model. These quantities will all depend on $\widehat{\mathbf{P}}$, $\widehat{\mathbf{T}}$, $\breve{\mathbf{P}}$ and $\breve{\mathbf{T}}$.

We partition $\widehat{\mathbf{P}}$, $\widehat{\mathbf{T}}$ and $\breve{\mathbf{P}}$, $\breve{\mathbf{T}}$ according to $\widehat{\mathcal{S}} \cup \mathcal{S}_-$ and $\breve{\mathcal{S}} \cup \mathcal{S}_+$, respectively:

$$\widehat{\mathbf{P}} = \left[\widehat{\mathbf{P}}_{+0} \,\vdots\, \widehat{\mathbf{P}}_{+-}\right],$$
$$\breve{\mathbf{P}} = \left[\breve{\mathbf{P}}_{-0} \,\vdots\, \breve{\mathbf{P}}_{-+}\right],$$
$$\widehat{\mathbf{T}} = \left[\widehat{\mathbf{T}}_{00} \,\vdots\, \widehat{\mathbf{T}}_{0-}\right],$$
$$\breve{\mathbf{T}} = \left[\breve{\mathbf{T}}_{00} \,\vdots\, \breve{\mathbf{T}}_{0+}\right].$$

Now, define the matrices

$$\mathbf{\Lambda}(s) = \widehat{\mathbf{P}}_{+0}\left(s\mathbf{I} - \widehat{\mathbf{T}}_{00}\right)^{-1}\widehat{\mathbf{T}}_{0-} + \widehat{\mathbf{P}}_{+-},$$
$$\mathbf{V}(s) = \breve{\mathbf{P}}_{-0}\left(s\mathbf{I} - \breve{\mathbf{T}}_{00}\right)^{-1}\breve{\mathbf{T}}_{0+} + \breve{\mathbf{P}}_{-+}.$$

$\mathbf{\Lambda}(s)$ and $\mathbf{V}(s)$ are the LSTs of the time spent during a visit to the upper and lower boundary, respectively. For instance, $[\mathbf{\Lambda}(s)]_{i,j}$ is the LST of the time spent at the upper boundary b, given the process hits b in phase $i \in \mathcal{S}_+$, and leaves the boundary in phase $j \in \mathcal{S}_-$. Likewise, $[\mathbf{V}(s)]_{i,j}$ is the LST of the time spent at the lower boundary 0, given the process hits 0 in phase $i \in \mathcal{S}_-$, and leaves the boundary in phase $j \in \mathcal{S}_+$.

Next, we define the matrices ${}_z^b\widehat{\mathbf{W}}(s)$ and ${}_0^z\breve{\mathbf{W}}(s)$ as follows:

$${}_z^b\widehat{\mathbf{W}}(s) = \mathbf{\Lambda}(s)\left(\mathbf{I} - {}_z^b\widehat{\mathbf{H}}(s;b)\mathbf{\Lambda}(s)\right)^{-1},$$
$${}_0^z\breve{\mathbf{W}}(s) = \mathbf{V}(s)\left(\mathbf{I} - {}_0^z\widehat{\mathbf{G}}(s;0)\mathbf{V}(s)\right)^{-1}.$$

These matrices are the respective LSTs of the time spent during a path which begins and ends at the upper or lower boundary, but is allowed to leave and come back to that boundary as long as the path does not hit the taboo level z.

For instance, $[{}_z^b\widehat{\mathbf{W}}(s)]_{i,j}$ is the LST of the time spent in the interval $(z,b]$, given that the process has just hit the upper boundary b in a phase $i \in \mathcal{S}_+$ and eventually leaves the boundary, after any number of journeys in the interval $(z,b]$, in a phase $j \in \mathcal{S}_-$. These matrices provide all that is necessary to construct the analogues to $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$.

We now introduce the important quantities ${}^{b}_{z}\widetilde{\mathbf{\Psi}}(s)$ and ${}^{z}_{0}\widetilde{\mathbf{\Xi}}(s)$. If we define the first passage time $\widetilde{\theta}(z) = \inf\{t > 0 : \widetilde{M}(t) = z\}$, then for $i \in \mathcal{S}_+$ and $j \in \mathcal{S}_-$,

$$\left[{}^{b}_{z}\widetilde{\mathbf{\Psi}}(s)\right]_{i,j} = \mathbb{E}\left[e^{-s\widetilde{\theta}(z)}\mathbb{I}\left(\widetilde{\theta}(z) < \infty, \widetilde{\varphi}(\widetilde{\theta}(z)) = j\right) \;\middle|\; \widetilde{M}(0) = z, \widetilde{\varphi}(0) = i\right],$$
$$\left[{}^{z}_{0}\widetilde{\mathbf{\Xi}}(s)\right]_{j,i} = \mathbb{E}\left[e^{-s\widetilde{\theta}(z)}\mathbb{I}\left(\widetilde{\theta}(z) < \infty, \widetilde{\varphi}(\widetilde{\theta}(z)) = i\right) \;\middle|\; \widetilde{M}(0) = z, \widetilde{\varphi}(0) = j\right].$$

These are the LSTs of the time taken to return to the initial level of z, given the process starts in an upwards phase and ends in a downwards phase (${}^{b}_{z}\widetilde{\mathbf{\Psi}}(s)$) or vice versa (${}^{z}_{0}\widetilde{\mathbf{\Xi}}(s)$). Theorem 3 in Bean *et al.* [13] states that

$${}^{b}_{z}\widetilde{\mathbf{\Psi}}(s) = {}^{b}_{z}\widehat{\mathbf{G}}_{+-}(s;z) + {}^{b}_{z}\widehat{\mathbf{H}}_{++}(s;z)\,{}^{b}_{z}\widehat{\mathbf{W}}(s)\,{}^{b}_{z}\widehat{\mathbf{G}}_{--}(s;b), \tag{2.26}$$
$${}^{z}_{0}\widetilde{\mathbf{\Xi}}(s) = {}^{z}_{0}\widehat{\mathbf{H}}_{-+}(s;z) + {}^{z}_{0}\widehat{\mathbf{G}}_{--}(s;z)\,{}^{z}_{0}\widecheck{\mathbf{W}}(s)\,{}^{z}_{0}\widehat{\mathbf{H}}_{++}(s;0). \tag{2.27}$$

This time, each quantity *does* depend on the starting level z, since this will determine the interval in which the path must lie. In both expressions, one can interpret the first element of the sum as representing all paths which return to the initial level without hitting the boundary (i.e. in taboo of the boundary), and the second element as representing all paths that hit the upper or lower boundary, with the ${}^{b}_{z}\widehat{\mathbf{W}}(s)$ or ${}^{z}_{0}\widecheck{\mathbf{W}}(s)$ terms accounting for the behaviour that ensues once the process hits the upper or lower boundary, respectively—except for the sub-path which returns to level z, which is captured by ${}^{b}_{z}\widehat{\mathbf{G}}_{--}(s;b)$ or ${}^{z}_{0}\widehat{\mathbf{H}}_{++}(s;0)$.

Finally, we once again use the first passage time $\widetilde{\theta}(z)$ to define the quantities ${}^{b}_{x}\widetilde{\mathbf{G}}(s;z)$ and ${}^{y}_{0}\widetilde{\mathbf{H}}(s;z)$ as follows: for $0 \le x < z < y \le b$ and $i, j \in \mathcal{S}_+ \cup \mathcal{S}_-$,

$$\left[{}^{b}_{x}\widetilde{\mathbf{G}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\widetilde{\theta}(x)}\mathbb{I}\left(\widetilde{\theta}(x) < \infty, \widetilde{\varphi}(\widetilde{\theta}(z)) = j\right) \;\middle|\; \widetilde{M}(0) = z, \widetilde{\varphi}(0) = i\right],$$
$$\left[{}^{y}_{0}\widetilde{\mathbf{H}}(s;z)\right]_{i,j} = \mathbb{E}\left[e^{-s\widetilde{\theta}(y)}\mathbb{I}\left(\widetilde{\theta}(y) < \infty, \widetilde{\varphi}(\widetilde{\theta}(z)) = j\right) \;\middle|\; \widetilde{M}(0) = z, \widetilde{\varphi}(0) = i\right].$$

When $x = z$ or $z = y$, we can use the same limiting arguments as we use to calculate ${}^{y}_{x}\widehat{\mathbf{G}}(s;x)$, ${}^{y}_{x}\widehat{\mathbf{G}}(s;y)$, ${}^{y}_{x}\widehat{\mathbf{H}}(s;x)$ and ${}^{y}_{x}\widehat{\mathbf{H}}(s;y)$.

${}^{b}_{x}\widetilde{\mathbf{G}}(s;z)$ and ${}^{y}_{0}\widetilde{\mathbf{H}}(s;z)$ are the LSTs of the time taken to respectively drain or fill the fluid buffer in the bounded model. ${}^{b}_{x}\widetilde{\mathbf{G}}(s;z)$ corresponds to draining the buffer from z to x on the interval $[x, b]$, and ${}^{y}_{0}\widetilde{\mathbf{H}}(s;z)$ corresponds to filling the buffer from z to y on the interval $[0, y]$.

Once again, these quantities can be partitioned according to $\mathcal{S}_+ \cup \mathcal{S}_-$. That is,

$${}^{b}_{x}\widetilde{\mathbf{G}}(s;z) = \begin{bmatrix} \mathbf{0} & {}^{b}_{x}\widetilde{\mathbf{G}}_{+-}(s;z) \\ \mathbf{0} & {}^{b}_{x}\widetilde{\mathbf{G}}_{--}(s;z) \end{bmatrix}, \quad {}^{y}_{0}\widetilde{\mathbf{H}}(s;z) = \begin{bmatrix} {}^{y}_{0}\widetilde{\mathbf{H}}_{++}(s;z) & \mathbf{0} \\ {}^{y}_{0}\widetilde{\mathbf{H}}_{-+}(s;z) & \mathbf{0} \end{bmatrix}.$$

Theorem 4 from Bean *et al.* [13] states that for $0 \leq x \leq z \leq y \leq b$,

$$
{}_{x}^{b}\widetilde{\mathbf{G}}_{--}(s;z) = \begin{cases} \left(\mathbf{I} - {}_{x}^{z}\widehat{\mathbf{H}}_{-+}(s;z)\,{}_{z}^{b}\widetilde{\mathbf{\Psi}}(s)\right)^{-1} {}_{x}^{b}\widehat{\mathbf{G}}_{--}(s;z), & x < z, \\ \mathbf{I} & x = z, \end{cases} \tag{2.28}
$$

$$
{}_{0}^{y}\widetilde{\mathbf{H}}_{++}(s;z) = \begin{cases} \left(\mathbf{I} - {}_{z}^{y}\widehat{\mathbf{G}}_{+-}(s;z)\,{}_{0}^{z}\widetilde{\mathbf{\Xi}}(s)\right)^{-1} {}_{0}^{y}\widehat{\mathbf{H}}_{++}(s;z) & z < y, \\ \mathbf{I} & z = y. \end{cases} \tag{2.29}
$$

The matrix inverses in each of these expressions account for any number of journeys back to the starting level z. Subsequently, for $0 \leq x \leq z \leq y \leq b$,

$$
{}_{x}^{b}\widetilde{\mathbf{G}}_{+-}(s;z) = {}_{z}^{b}\widetilde{\mathbf{\Psi}}(s)\,{}_{x}^{b}\widetilde{\mathbf{G}}_{--}(s;z), \tag{2.30}
$$

$$
{}_{0}^{y}\widetilde{\mathbf{H}}_{-+}(s;z) = {}_{0}^{z}\widetilde{\mathbf{\Xi}}(s)\,{}_{0}^{y}\widetilde{\mathbf{H}}_{++}(s;z). \tag{2.31}
$$

${}_{z}^{b}\widetilde{\mathbf{\Psi}}(s)$ and ${}_{0}^{z}\widetilde{\mathbf{\Xi}}(s)$ are almost identical to $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$, with the addition of the upper and lower boundary, respectively. Similarly, ${}_{x}^{b}\widetilde{\mathbf{G}}(s;z)$ and ${}_{0}^{y}\widetilde{\mathbf{H}}(s;z)$ are almost identical to ${}_{x}^{y}\widehat{\mathbf{G}}(s;z)$ and ${}_{x}^{y}\widetilde{\mathbf{H}}(s;z)$, except the taboo level is replaced with the upper and lower boundary, respectively. Rather than the LSTs ignoring all paths which hit this special level, these paths are now included—except in the bounded model, such paths now have special behaviour at the boundary.

Onwards and upwards (in $i \in \mathcal{S}_{+}$)

This ends the background necessary to understand the model which this thesis will revolve around. It is particularly important to understand the LSTs ${}_{z}^{b}\widetilde{\mathbf{\Psi}}(s)$, ${}_{0}^{z}\widetilde{\mathbf{\Xi}}(s)$, ${}_{x}^{b}\widetilde{\mathbf{G}}(s;z)$ and ${}_{0}^{y}\widetilde{\mathbf{H}}(s;z)$ and their interpretations, as these quantities hold particular significance in what is to come. In any case, we now move onwards and upwards (in an upwards phase) to the model for coral-algal symbiosis.

Chapter 3

A Model for Coral-Algal Symbiosis

There are many existing models for coral bleaching, but rarely do these models account for the random factors associated with coral-algal symbiosis, of which there is an abundance. In this chapter, we introduce a stochastic fluid model to model the mechanics of coral-algal symbiosis in a stochastic manner. We motivate and state the assumptions of the model, inspired by the features of coral-algal symbiosis covered in Section 2.1. We specify the model mathematically, and define the process of coral mortality. We also specify a set of canonical parameters which we use to test the model in future chapters. Finally, we derive an important quantity in the model, called the time to mortality, which will reappear many times throughout this thesis.

3.1 Motivation

In Section 2.2, we looked at existing models for coral bleaching, and saw that many existing models focus on the large-scale effects of bleaching. There are also some models for the small-scale dynamics of coral-algal symbiosis, although these models are in the minority. A common factor among such models is that they are *deterministic*.

There are many factors which determine the severity of bleaching events, as well as the degree of recovery from such events: the species of coral, of which there is no lack of diversity on the GBR; the genus of *Symbiodiniaceae*; environmental factors such as temperature and UV; ecological factors such as water quality and geographical location.

It is therefore appropriate to formulate a *stochastic* model for coral bleaching, which is the aim of this chapter.

3.2 The model

The model we introduce here originates from an unpublished 2009 paper titled *A Stochastic Fluid Model of the Adaptive Bleaching Hypothesis on the Great Barrier Reef* [42]. As the name suggests, the model intended to shed light on the Adaptive Bleaching Hypothesis (ABH), introduced in Section 2.1.4. The ABH has since fallen out of favour, but the model itself is still useful. The majority of this chapter follows directly from this work, with the addition of some extra (and updated) biological understanding, and a different set of parameters.

Throughout this section, we use the numbers 1-7 to refer to the following modelling features and assumptions (which are themselves related).

3.2.1 Features

Our model attempts to capture the following key features of coral-algal symbiosis and coral bleaching. The biology behind these features is fully detailed in Section 2.1.

1. The number of *Symbiodiniaceae* in a coral colony is approximately one million algal cells per square centimetre of coral tissue.

2. Coral can uptake *Symbiodiniaceae* from the environment via symbiont switching, if its entire symbiont population dies (note: this is typically within juvenile corals, but we assume this is possible in our coral host for convenience).

3. Coral can form symbiotic relationships with multiple genera of *Symbiodiniaceae* at the same time (the coral can harbour multiple genera), and over time (the coral can form relationships with a new genus). On the GBR, the most prevalent genera are *S. Cladocopium* and *S. Durusdinium* which we henceforth refer to as *genus C* and *genus D*, respectively.

4. During stressful conditions, a coral host may bleach, expelling some or all of its resident symbiont population. Stressful conditions are a result of various environmental and ecological factors, and vary in duration.

5. Genus C has a higher rate of photochemical efficiency than genus D.

6. Genus D is more thermally tolerant than genus C.

7. Coral stores excess energy received from its symbiont in the form of lipids. When the symbiont population is expelled during a bleaching event, the coral may be required to sustain itself on these energy reserves. If bleaching persists for too long and the coral runs out of stored energy, it dies.

3.2.2 Assumptions

In order to capture the biological features of coral-algal symbiosis into a model, we make the following assumptions:

1. We consider an arbitrary coral 'unit', and represent the population of *Symbiodiniaceae* in this unit as a continuous density between 0 and 1. This is referred to as the *algal density*.

2. If the algal density reaches 0, the density can recover via symbiont switching.

3. Genus C and D may be hosted within the coral simultaneously, but we only consider the 'dominant' symbiont genus (the genus with the largest population). The dominant genus can change under certain circumstances.

4. A bleaching event can occur at any time, during which the algal density rapidly decreases at a constant rate; the event lasts for a random duration of time. When a bleaching event ends, either genera can become the dominant genus. Genus D is more likely to become dominant after a bleaching event. When a bleaching event is not occurring, the algal density grows at a constant rate that is much smaller than the bleaching rate.

5. Genus C has a lower growth rate than genus D, and the coral host is more likely to bleach when genus C is dominant. The dominant genus can switch from D to C when the density is 1 (but not from C to D).

6. Genus D has a higher growth rate than genus C, and the coral host is less likely to bleach when genus D is dominant. The dominant genus can switch from C to D when the density is less than 1 (but not from D to C).

7. There is a critical algal density at which the coral receives exactly enough energy from its symbiont population to survive. Above this density, the coral can store excess energy in the form of lipids; below this density, it must rely on stored lipids. If the coral spends too long below this threshold, it dies.

3.2.3 Model specification

Here, we describe a stochastic fluid flow model for coral-algal symbiosis. The model behaves identically to $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0} = \{(\widetilde{M}(t), \widetilde{\varphi}(t)) \in [0, b] \times \mathcal{S} \cup \widehat{\mathcal{S}} \cup \check{\mathcal{S}} : t \geq 0\}$ from Section 2.4.2, but with the addition of a mortality process. Here, we specify the requisite fluid matrices and the process of mortality.

This model is referred to as the *basic model*.

In the basic model, the level $\widetilde{M}(t)$ represents the algal density; thus, $b = 1$, and the level evolves in the interval $[0, 1]$. The phase $\widetilde{\varphi}(t)$ represents the dominant genus of *Symbiodiniaceae* and whether algal growth (increasing density) or bleaching (decreasing density) is occurring.

To specify the basic model, we must specify the state spaces $\mathcal{S}$, $\widehat{\mathcal{S}}$ and $\check{\mathcal{S}}$, as well as the matrices $\mathbf{T}$, $\mathbf{C}$, $\widehat{\mathbf{P}}$, $\widehat{\mathbf{T}}$, $\check{\mathbf{P}}$ and $\check{\mathbf{T}}$ (we specify $\mathbf{T}$, $\widehat{\mathbf{T}}$, and $\check{\mathbf{T}}$ last).

Recall that we may partition $\mathcal{S}$ according to whether the fluid level is increasing or decreasing in each state; that is, $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_-$. We define $\mathcal{S}_+ = \{C_+, D_+\}$ and $\mathcal{S}_- = \{C_-, D_-\}$. Genus C is dominant in the phases C_+ and C_-; likewise, genus D is dominant in the phases D_+ and D_-. Phases in $\mathcal{S}_+$ represent phases where the algal density is increasing, corresponding to algal growth. Phases in $\mathcal{S}_-$ represent phases where the algal density is decreasing, corresponding to a bleaching event.

We denote the fluid input/output rates as c_+^C, c_+^D, c_-^C and c_-^D. That is, when genus C is dominant during growth, the fluid buffer fills at a rate of c_+^C; when genus C is dominant during a bleaching event, the fluid buffer drains at a rate of c_-^C; and similarly for genus D. We define

$$\mathbf{C} = \begin{bmatrix} c_+^C & 0 & 0 & 0 \\ 0 & c_+^D & 0 & 0 \\ 0 & 0 & c_-^C & 0 \\ 0 & 0 & 0 & c_-^D \end{bmatrix}.$$

At the upper boundary $b = 1$, there are only two states, representing which genus is dominant, so we define $\widehat{\mathcal{S}} = \{\widehat{C}, \widehat{D}\}$. Recall that in the bounded fluid model, when the level hits the upper boundary, it enters $\widehat{\mathcal{S}} \cup \mathcal{S}_-$ according to $\widehat{\mathbf{P}}$. In our case, the process never reflects off the boundary: therefore, we define

$$\widehat{\mathbf{P}} = \left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right].$$

At the lower boundary 0, there is only one state, since there is no dominant genus if the algal density is zero, so we define $\check{\mathcal{S}} = \{\check{0}\}$. Recall that in the bounded fluid model, when the level hits the lower boundary, it enters $\check{\mathcal{S}} \cup \mathcal{S}_+$ according to $\check{\mathbf{P}}$. Similarly to the upper boundary, the process never reflects off the lower boundary: therefore, we define

$$\check{\mathbf{P}} = \left[\begin{array}{c:cc} 1 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right].$$

Infinitesimal generators

For positive constants α, β, γ, δ, η and $r \geq 1$, we define $\mathbf{T}$, $\widehat{\mathbf{T}}$ and $\breve{\mathbf{T}}$ as follows:

$$\mathbf{T} = \begin{bmatrix} -(\beta + r\alpha) & \beta & r\alpha & 0 \\ 0 & -\alpha & 0 & \alpha \\ \gamma & \delta & -(\gamma+\delta) & 0 \\ \gamma & \delta & 0 & -(\gamma+\delta) \end{bmatrix},$$

$$\widehat{\mathbf{T}} = \left[\begin{array}{cc:cc} -r\alpha & 0 & r\alpha & 0 \\ \eta & -(\alpha+\eta) & 0 & \alpha \end{array}\right],$$

$$\breve{\mathbf{T}} = \left[\begin{array}{c:cc} -(\gamma+\delta) & \gamma & \delta \end{array}\right],$$

which may be interpreted as follows:

- When $\widetilde{M}(t) \in (0,1)$ and $\widetilde{\varphi}(t) = C_+$ bleaching events (transitions into C_-) occur at rate $r\alpha$, and genus D becomes dominant (transition into D_+) at rate β. When $\widetilde{M}(t) \in (0,1)$ and $\widetilde{\varphi}(t) = D_+$, bleaching events (transitions into D_-) occur at rate α.

- When $\widetilde{M}(t) \in (0,1)$ and $\widetilde{\varphi}(t) = C_-$ or D_-, the coral recovers from bleaching at rate $\gamma + \delta$. When bleaching ends, genus C becomes dominant (transitions into C_+) with probability $\gamma/(\gamma+\delta)$, and genus D becomes dominant (transitions into D_+) with probability $\delta/(\gamma+\delta)$.

- When $\widetilde{M}(t) = 1$, bleaching events occur at rates $r\alpha$ and α from $\widehat{C}$ and $\widehat{D}$ into C_- and D_-, respectively. If $\widetilde{\varphi}(t) = \widehat{D}$, genus C becomes dominant (transitions into $\widehat{C}$) at rate η.

- When $\widetilde{M}(t) = 0$, the coral uptakes *Symbiodiniaceae* at rate $\gamma + \delta$. When this occurs, genus C becomes dominant (transitions into C_+) with probability $\gamma/(\gamma+\delta)$, and genus D becomes dominant (transitions into D_+) with probability $\delta/(\gamma+\delta)$.

Several assumptions have been made here. First, we have defined the bleaching rate for genus C as $r\alpha$. This is predominantly to address Assumption 5 and Assumption 6; provided $r > 1$, the bleaching rate for genus C is forced to be higher than the rate for genus D. We may also interpret r as the advantage of genus D over genus C due to the enhanced thermal tolerance of genus D.

Second, we assume that the bleaching behaviour at the upper boundary is the same as when the density is less than 1. This is realistic biologically and makes our model easier to parameterise.

Last, we assume that the uptake process from $\breve{0}$ is the same as the recovery process from $\mathcal{S}_-$. That is, the rate of uptake is the same as the rate of recovery, and the probabilities of each genus becoming dominant are the same. This is to simplify the model and reduce the number of parameters required.

The mortality process

To define the process of mortality, we use Assumption 7, which we explicitly formulate here. First, we define the coral's energy process as a process $E(t)$. By Assumption 1, it is reasonable to assume that it has some constant metabolic energy requirement R, which is the rate of energy consumption over time. Then we may say that

$$\frac{dE(t)}{dt} = \widetilde{M}(t)P_{\widetilde{\varphi}(t)} - R, \tag{3.1}$$

where $P_{\widetilde{\varphi}(t)}$ is the (constant) energy production rate of the coral's symbiont population, when the phase is $\widetilde{\varphi}(t)$. This depends on the current dominant genus of *Symbiodiniaceae* only, and not whether growth or bleaching is occurring.

Then $dE(t)/dt = 0$ when $\widetilde{M}(t)P_{\widetilde{\varphi}(t)} = R$. Define $\zeta \in (0,1)$ as the level such that $\zeta P_{\widetilde{\varphi}(t)} = R$. That is, ζ is the algal density such that the energy provided to the coral host by its symbiont is exactly equal to the coral's energy requirement. We refer to ζ as the *critical threshold*. For simplicity, we assume for now that $P_{\widetilde{\varphi}(t)}$ is independent of the phase $\widetilde{\varphi}(t)$, and thus ζ is a constant model parameter.

Feature 7 helps us describe the behaviour of the energy process above and below the threshold. When $\widetilde{M}(t) > \zeta$, the energy being produced exceeds the coral's energy requirement, and the coral stores this excess energy in the form of lipids. When $\widetilde{M}(t) < \zeta$, the coral is in energy deficit, and must draw upon these stored lipids. The big question is: when does the coral run out of stored energy?

We answer this question using an approximation: we assume that the coral runs out of stored energy when the algal density spends a fixed time τ below the critical threshold ζ, on a *single visit* below the threshold. That is, when the algal density goes below ζ, it must return to ζ within some time $t < \tau$. If it does, then the process continues, and we say that the coral has returned to full energy. If it remains below ζ for some time $t \geq \tau$, then the coral dies.

This approximation does not account for the recovery process of the coral above ζ. That is, if the algal density returns to ζ, the behaviour of the process below ζ, which represents a net energy loss, is forgotten entirely. However, it is reasonable to assume such behaviour: in reality, bleaching events are rare in comparison to periods of growth, and so under our model, the coral is likely to

Disclaimer

The word *canonical* is important! We attempt to make our parameter choices as sensible as possible, and to represent the true process of coral bleaching on the GBR as closely as possible within the constraints of the model and the limitations on available data. We do not claim that these parameters accurately represent coral or coral bleaching on the GBR. In this thesis, they will mainly be used for testing and analysing the fluid model(s) we construct, in addition to explaining what our results mean *if the parameters of our model are correct.*

A better set of parameters could be obtained via experiments, but this is outside the scope of this thesis.

recover fully if the algal density returns to ζ. Moreover, the 'memoryless' nature of this assumption allows us to use the fluid model framework to model coral-algal symbiosis, since we are able to obtain the distribution of time spend below a fixed threshold ζ, as we shall see in Section 3.3.

In fairness, our model would benefit if the behaviour under the critical threshold contributed to future bleaching events. If this were so, then the model would contain some sort of bleaching history; this may or may not be realistic, but the additional feature would make the model more versatile.

The parameters ζ and τ are representative of the coral's survivability. Intuitively, ζ represents the coral's resistance to bleaching events—a lower threshold means that the severity of bleaching must be higher in order for the coral to begin to lose energy. The parameter τ represents the coral's relative energy storage—a higher value means that the coral can survive for longer on its stored energy.

3.2.4 Canonical model parameters

We now proceed to choose values for the aforementioned parameters. We refer to these parameters as the *canonical parameters*, and the resulting model as the *canonical model*. It is important to understand the use of the word 'canonical' here, which implies that these parameters are chosen to represent coral on the GBR as closely as possible, but may not achieve that goal entirely (see Disclaimer).

Note that these parameter differ from those in Helfgott *et al.* [42].

The parameters of the canonical model are given in Table 3.1, where rates are given in density per day and τ is in days. We proceed to explain the reasoning behind these choices—in roughly descending order, with respect to Table 3.1.

Parameter	Interpretation	Value
c_+^C	Algal growth rate during ambient conditions, when genus C is dominant.	$\frac{1}{30}$
c_+^D	Algal growth rate during ambient conditions, when genus D is dominant.	$\frac{2}{30}$
$c_-^C = c_-^D$	Rate of symbiont expulsion during a bleaching event.	-1
α	Bleaching event rate when the dominant genus is genus D. The mean time before an event occurs is $1/\alpha$.	$\frac{1}{5\times365}$
r	Disadvantage of genus C in relation to genus D, due to lack of thermal tolerance. Multiplies the bleaching rate of genus D to obtain the rate for genus C.	2
γ	Rate at which bleaching events end, and genus C becomes dominant.	$\frac{1}{2}$
δ	Rate at which bleaching events end, and genus D becomes dominant.	1
β	Rate at which genus C is succeeded by genus D in the interval $(0, 1)$.	$\frac{10}{5\times365}$
η	Rate at which genus D is succeeded by genus C at the upper boundary.	$\frac{10}{5\times365}$
ζ	Critical algal density. When the density is below this threshold, the coral is in energy deficit.	0.5
τ	Days that the algal density can remain below the critical threshold on a single visit before mortality occurs.	14

Table 3.1: Parameters in the canonical model, and their interpretations: separated into fluid input/output rates (top), parameters related to transition rates (middle) and mortality parameters (bottom). Rate parameters are given in density per day.

- In general, algal growth rates depend on the current population, and are more likely to be logistic than constant [54]. Furthermore, algal growth rates can be highly variable, and experimental growth rates are often characterised by mass and cell proliferation, as opposed to density, making it difficult to choose values for c_+^C and c_+^D. We choose values of 1/30 and 2/30 for the growth rates of genus C and D, respectively, with the interpretation that a fully bleached coral harbouring genus C can recover to maximum algal density in a month, and in half that time when genus D is dominant.

- On the scale of days, symbiont expulsion due to bleaching is almost immediate, and thus the rate of symbiont expulsion is a difficult rate to choose. Rather, we calibrate the expulsion rate and the rate of bleaching events ending (recovery rate). We choose the value of $c_-^C = c_-^D = -1$ for the expulsion rate, and a value of $\gamma + \delta = 1.5$ for the recovery rate. Thus, the density expelled during bleaching has an exponential distribution with rate 1.5, with the caveat that any density expelled that exceeds 1 is added to a point mass at 1, due to the restriction imposed by the lower boundary.

  This distribution is chosen such that there is a high probability of the density ending in the interval $(0, \zeta)$. As a result, the process is more likely to create bleaching events extreme enough to end up below the critical threshold, but not too extreme, as to always hit the lower boundary.

  We set $\gamma = 0.5$ and $\delta = 1$ such that after a bleaching event genus D is twice as likely to become dominant, as a result of Assumption 6.

- For the bleaching event rate, we want to capture the idea that bleaching events are rare, but if severe enough, can kill a coral colony. A value of $\alpha = 1/(5 \times 365)$ suits this purpose, and has the interpretation that bleaching events occur, on average, every five years. We choose a value of $r = 2$, such that events happen twice as often when genus C is dominant, since genus C is more susceptible to bleaching compared to genus D, satisfying Assumption 5.

- The succession rates (from genus C to D, and vice-versa) are chosen to be ten times the bleaching rate α, so $\beta = \eta = 10\alpha$. As a result, succession is quite likely to occur between bleaching events, although not guaranteed.

- The parameter τ is an abstraction of the energy reserves of the coral host, and our choice for this parameter is somewhat arbitrary. We choose a value of $\tau = 14$ which gives the interpretation that if the coral cannot recover within two weeks, after a bleaching event which leaves it in energy deficit, then it will die.

- The parameter ζ is an abstraction of the coral's resilience to bleaching, and once again our choice is rather arbitrary. Moreover, the choice of ζ depends on many other parameters such as $c_-^C = c_-^D$ and $\gamma + \delta$. In accordance with the aforementioned parameter choices, we choose a value of $\zeta = 0.5$, since this value attributes a high probability towards bleaching events which begin at the upper boundary $b = 1$ and end up in the interval $(0, \zeta)$. This way, bleaching events are severe enough to frequently bring the coral into energy deficit, but not so severe as to frequently drain the algal density to 0.

We are primarily interested in analysing the mathematical and numerical properties of this model, and so we proceed with these canonical parameters.

3.2.5 Assumptions revisited

With our model fully specified, we revisit our assumptions in an effort to justify them, and explain why they are needed to mathematically model coral-algal symbiosis with a stochastic fluid model.

1. We assume a fixed coral 'unit' to abstract the problem away from a coral colony or reef. Although corals may not be homogeneous in the size of their polyps and the algal density within them, it is reasonable to assume that we can represent the algal growth of the coral in this unit. Due to the large algal population within each polyp, it is reasonable to assume that the population is a continuous density.

2. To account for cases where a coral is fully bleached (the algal density is 0) and recovers, we allow the density to recover from 0. The mechanism of symbiont switching is one justification for this, although, as mentioned, this usually only happens in juvenile corals. A more reasonable justification may be that the algal density never really reaches 0, and that in our model, level 0 means a small algal population that is still able to recover.

3. Ideally, we would track the algal densities of each algal genus over time; however, due to the restrictions of mathematical models, this is either very difficult or impossible. Instead, we record the dominant genus only, as an approximation. Furthermore, in the stochastic fluid model framework, we may add more phases to the phase process to add more complex behaviour, but we only consider the dominant genus throughout this thesis. The possibility of more complicated genus dynamics is addressed in Chapter 7.

4. Rather than modelling the conditions which may lead to bleaching—such as coral species, temperature, UV, water quality geographical location, and so on—we simply model bleaching events as a random process. This abstracts the notion of bleaching away from the various complex factors which may cause bleaching events, meaning we do not need to assume anything about such conditions. We are able to control the bleaching process directly via parameters such as α, γ, δ, c_-^C and c_-^D.

   We assume that the algal growth rate is constant for simplicity, but logistic growth may be more sensible. This is addressed in more detail in Chapter 7.

5. We address the assumption of algal growth rates by specifying the parameters c_+^C and c_+^D. Although the canonical choices may not be perfect, the growth rates can be different, and we are allowed a fair degree of control over these rates. We incorporate the fact that the dominant genus can switch from C to D when the algal density is less than 1 via the parameter β, and can switch from D to C at when the algal density is 1 via the parameter η. The parameters γ and δ allow for different bleaching rates.

   Note: we use Feature 5 to justify the assumption that genus D has a higher growth rate and can succeed genus C when the algal density is less than 1. The enhanced photochemical efficiency of genus C justifies the ability to switch from genus D to genus C when the algal density is 1.

6. See 5.

7. Assumption 7 is key to the entire model, since it defines the process of mortality. To use this assumption to define a model, we must make additional assumptions. In the basic model (which is the basis for Chapter 4), this means assuming that the process can only spend a fixed time τ below the critical threshold. In later chapters, we make different assumptions to define the process of mortality, in an attempt to improve on the basic model.

3.3 The time to mortality

Within our model, we defined coral mortality to be the time at which the coral has spent too long below the critical threshold ζ and dies of energy starvation. The time at which mortality occurs is of great interest to us, since it tells us about the resilience of coral to bleaching.

We define the *time to mortality* (TTM) to be the exact moment that the coral dies under our model. Using stochastic fluid models, we can derive the *distribution*

of the TTM. The distribution of the TTM provides much more information than a simple measure such as the mean time of mortality. In this section, we obtain the Laplace-Stieltjes transform (LST) of the TTM; we can then invert this transform to recover the distribution of the TTM.

The first crucial step towards obtaining the LST of the TTM is finding the LST of the time spent below the critical threshold ζ.

3.3.1 Time spent below the critical threshold

Recall from Section 2.4.2 that ${}_0^z\widetilde{\boldsymbol{\Xi}}(s)$ is the matrix LST of the time spent below level z, before returning to z, in the bounded process $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$. We define $\boldsymbol{\xi}(t)$ as the (matrix) probability density function (PDF) associated with ${}_0^\zeta\widetilde{\boldsymbol{\Xi}}(s)$. Therefore, $\boldsymbol{\xi}(t)$ is the inverse LST of ${}_0^\zeta\widetilde{\boldsymbol{\Xi}}(s)$, and

$$ {}_0^z\widetilde{\boldsymbol{\Xi}}(s) = \int_0^\infty e^{-su}\boldsymbol{\xi}(u)\,du. $$

We would like to obtain the LST of the time spent below ζ, such that the process returns to ζ before τ has elapsed. This (matrix) LST is

$$ \left[\int_0^\tau e^{-su}\boldsymbol{\xi}(u)\,du\right]_{j,i} = \mathbb{E}\left[e^{-s\widetilde{\theta}(\zeta)}\mathbb{I}\left(\widetilde{\theta}(\zeta)<\tau, \widetilde{\varphi}(\widetilde{\theta}(\zeta))=i\right) \,\middle|\, \widetilde{M}(0)=z, \widetilde{\varphi}(0)=j\right]. \tag{3.2} $$

Eq. (3.2) is the LST of $\boldsymbol{\xi}(t)$ censored at time τ. That is, we only consider values of $\boldsymbol{\xi}(u)$ where u is less than τ, corresponding to times where the process has returned to ζ before τ has elapsed.

Note that to obtain the LST in Eq. (3.2), we must perform a LST inversion. This is in addition to the LST inversion required to obtain the distribution of the TTM, which shall become very important in Chapter 4.

Also note that the PDF corresponding to Eq. (3.2) integrates[1] to a value less than 1 (as long as $\tau<\infty$) since there is a positive probability of mortality occurring. We say that such a distribution is *dishonest*; in comparison, the PDF of an *honest* distribution integrates[1] to 1.

[1]For matrix quantities such as the quantity in Eq. (3.2), we must consider all row sums of the integral. These values decide whether the matrix PDF is *honest* or *dishonest*.

3.3.2 Distribution of the time to mortality

Obtaining the distribution of the TTM relies heavily on the *convolution theorem*. This fundamental theorem underpins stochastic fluid model analysis, in that it allows us to obtain many time-dependent LSTs. The theorem is stated as follows.

Theorem 3.1. *Convolution theorem*

Let X and Y be independent random variables with LSTs $\widehat{F}_x(s)$ and $\widehat{F}_y(s)$, respectively. Then the LST of $X+Y$ is $\widehat{F}_x(s)\widehat{F}_y(s)$.

Using the convolution theorem, we can obtain the LST of the TTM. To do so, we consider a path in the basic model which leads to mortality beginning at the upper boundary $b = 1$ in a downwards phase $i \in \mathcal{S}_-$. This represents a perfectly healthy coral colony at the time of a bleaching event.

1. The path begins at $b = 1$. For mortality to occur, the level must reach ζ for the first time.

2. Once the algal density is at ζ, the process can make any number of down-up journeys below ζ which return to ζ, followed by up-down journeys above ζ which return to ζ, as long as each down-up journey is shorter than τ. This can happen $m \geq 0$ times.

3. After m such journeys, the process makes a single journey below ζ which exceeds τ in duration, and thus mortality occurs.

To calculate the LST of the TTM, we first derive the LSTs corresponding to each of the above sub-paths. The idea is that we can use the convolution theorem to combine these LSTs to obtain the LST of the TTM, since they are conditionally independent given the phase. On the following page, we provide an example of how this is possible with matrix LSTs. The LSTs of each sub-path are derived as follows:

1. The LST of the time taken, starting from level $b = 1$ in a downwards phase, to reach level ζ for the first time is ${}^{b}_{\zeta}\widetilde{\mathbf{G}}_{--}(s; b)$.

2. The LST of the time taken to make a down-up journey from ζ back to ζ, such that the process returns before τ has elapsed is the expression in Eq. (3.2).

   The LST of the time taken to make an up-down journey from ζ back to ζ is ${}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)$. Thus, the LST of the time taken to make exactly m such journeys is

$$\left[\int_0^\tau e^{-su}\boldsymbol{\xi}(u)\,du\,{}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)\right]^m.$$

Matrix convolution example

The convolution theorem is extremely helpful due to the ability to convert convolutions into products of LSTs. However, one might ask whether this is also possible for matrix LSTs? The answer is yes, due to conditional independence.

Consider $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$: in the basic model, these are 2×2 matrix LSTs. Let us define the product as $\widehat{\mathbf{\Phi}}(s) = \widehat{\mathbf{\Psi}}(s)\widehat{\mathbf{\Xi}}(s)$, which has entries:

$$\begin{aligned}
[\widehat{\mathbf{\Phi}}(s)]_{1,1} &= [\widehat{\mathbf{\Psi}}(s)]_{1,1}[\widehat{\mathbf{\Xi}}(s)]_{1,1} + [\widehat{\mathbf{\Psi}}(s)]_{1,2}[\widehat{\mathbf{\Xi}}(s)]_{2,1} \\
[\widehat{\mathbf{\Phi}}(s)]_{1,2} &= [\widehat{\mathbf{\Psi}}(s)]_{1,1}[\widehat{\mathbf{\Xi}}(s)]_{1,2} + [\widehat{\mathbf{\Psi}}(s)]_{1,2}[\widehat{\mathbf{\Xi}}(s)]_{2,2} \\
[\widehat{\mathbf{\Phi}}(s)]_{2,1} &= [\widehat{\mathbf{\Psi}}(s)]_{2,1}[\widehat{\mathbf{\Xi}}(s)]_{1,1} + [\widehat{\mathbf{\Psi}}(s)]_{2,2}[\widehat{\mathbf{\Xi}}(s)]_{2,1} \\
[\widehat{\mathbf{\Phi}}(s)]_{2,2} &= [\widehat{\mathbf{\Psi}}(s)]_{2,1}[\widehat{\mathbf{\Xi}}(s)]_{1,2} + [\widehat{\mathbf{\Psi}}(s)]_{2,2}[\widehat{\mathbf{\Xi}}(s)]_{2,2},
\end{aligned} \tag{3.3}$$

corresponding to the different start and end phases. Conditioned on these phases, the LST is scalar, and by the independence of the $[\widehat{\mathbf{\Psi}}(s)]_{i,k}$ and $[\widehat{\mathbf{\Xi}}(s)]_{k,j}$, for all i, j and k, the convolution theorem can be applied. The sums of LSTs are also well defined since they involve disjoint events (transitioning to phase 1 or 2).

The convolution theorem applies to matrix LSTs, since the matrix algebra conditions on the phase, allowing us to multiply independent scalar LSTs.

Thus, the LST of the time taken to make *any* number of such journeys is

$$\sum_{m=0}^{\infty}\left[\int_0^{\tau} e^{-su}\boldsymbol{\xi}(u)\,du\,{}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right]^m = \left(\mathbf{I}-\int_0^{\tau} e^{-su}\boldsymbol{\xi}(u)\,du\,{}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right)^{-1}.$$

This matrix inverse always exists: $\int_0^{\tau} e^{-su}\boldsymbol{\xi}(u)\,du$ is the LST of a dishonest distribution, and ${}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)$ is the LST of an honest distribution.[2] Thus, the product is the LST of a dishonest distribution.

3. On a journey below ζ which leads to mortality, exactly τ time elapses. The probability of returning *before* τ is

$$\int_0^{\tau}\boldsymbol{\xi}(u)\,du\,\mathbf{1}_+,$$

where $\mathbf{1}_+$ denotes a vector of ones of size s_+, since the end-phase does not interest us. Thus, the LST corresponding to this part of the journey is

$$e^{-s\tau}\left(\mathbf{1}_- - \int_0^{\tau}\boldsymbol{\xi}(u)\,du\,\mathbf{1}_+\right).$$

Finally, we can obtain the LST of the TTM.

Theorem 3.2. *The LST of the TTM under the basic model, given that the process starts in phase $i \in \mathcal{S}_-$ at level b, is*

$$\widehat{F}_{\mathrm{TTM}}(s) = \boldsymbol{e}_i^{\intercal}\,{}_{\zeta}^{b}\widetilde{\mathbf{G}}_{--}(s;b)\left(\mathbf{I}-\int_0^{\tau} e^{-su}\boldsymbol{\xi}(u)\,du\,{}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right)^{-1} e^{-s\tau}\left(\mathbf{1}_- - \int_0^{\tau}\boldsymbol{\xi}(u)\,du\,\mathbf{1}_+\right). \tag{3.4}$$

Proof. The result follows directly from the construction in the preceding section. Since, given the phase, each sub-path is conditionally independent, the convolution theorem holds, and we can simply multiply the LSTs corresponding to each sub-path to obtain the LST of the TTM.

Since the process begins in phase $i \in \mathcal{S}_-$, we pre-multiply by the unit vector $\boldsymbol{e}_i^{\intercal}$. ■

3.3.3 New directions

There are several drawbacks to the basic model: one of these is the approximation to the explicit energy process $E(t)$. Energy is the underlying driving force of

[2] The level process almost surely returns to ζ in a finite amount of time, since the process is bounded and consequently, level ζ is positive recurrent

coral mortality, and therefore paramount to modelling mortality. In this model, a big assumption is made in our definition of mortality: we do not account for the time for energy recovery above the threshold ζ, but simply assume that the coral instantaneously recovers its energy fully. This is reasonable enough, but we could improve our model by specifying an explicit energy process or bleaching history.

Moreover, we assume in Eq. (3.1) that $P_{\widetilde{\varphi}(t)}$ is the same for all phases, contradictory to the knowledge that different genera of *Symbiodiniaceae* produce energy at different rates. As a result, one of the key features that distinguishes genus C from genus D is not present in this model. This is particularly problematic since we are interested in answering questions about the differences between genus C and genus D, and this is one of the main distinctions between them.

Another drawback to the basic model is that τ is fixed. Previously, we stressed the importance of incorporating the random elements of coral bleaching; however, defining τ to be constant means that the rate of energy depletion and the energy buffer of the coral is always fixed. If τ were allowed to be a random variable, the mortality process in our model would be more flexible.

A simple way to incorporate this into the basic model is to redefine the second and third components of the LST of the TTM: for the second component, we could use a generic complementary CDF rather than the complementary CDF of a deterministic random variable with value τ; further related changes would be required in the third component.

This would address our τ-related concerns, but leave our other desires unfulfilled. We are therefore left with two paths to follow. On one hand, we have a working model and would like to explore some methods for obtaining the TTM under this model. On the other, we would like to address the need for an explicit energy process, which appears to require more sophisticated modelling techniques.

The following chapters will be dedicated to these two paths. We use the basic model to examine methods for obtaining the distribution of the time to mortality in Chapter 4, and then move on to exploring more sophisticated models in Chapter 5 and Chapter 6.

Chapter 4

The Time to Mortality

The time to mortality (TTM) is a proxy for the survivability of the coral host in our model for coral-algal symbiosis, and is a useful tool to provide insight into that process. In this chapter, we obtain the distribution for the TTM by numerically inverting the Laplace-Stieltjes transform (LST) obtained in Eq. (3.4). LST inversion methods can behave badly when applied to discontinuous functions, and so we investigate a time-shifting method to overcome these issues. We apply this time-shifting method to the TTM and discuss the results.

4.1 Introduction

In this section, we consider methods for inverting the LST of the TTM, denoted $\widehat{F}_{\mathrm{TTM}}(s)$, to obtain the inverse transform and hence evaluate $F_{\mathrm{TTM}}(t)$. We only consider the Euler [1] and CME [48, 49, 51] methods, both of which fit inside the *Abate-Whitt framework*.

Much of the background theory of Laplace transform inversion and the Abate-Whitt framework can be found in Section 2.3, including information about the Euler and CME methods. It is recommended to familiarise oneself with these concepts. It is worth noting that although the background covers the inversion of Laplace transforms, we will deal with *Laplace-Stieltjes* transforms in this section; however, the methods and methodology for LST inversion are exactly the same. It is also implicit that the LST inversion in this chapter is *numeric*, not *analytic*.

4.1.1 Methods

Inverting an LST under the Abate-Whitt framework can be broken into three steps: constructing a set of nodes (β_k) and weights (η_k), evaluating the LST at these nodes, and adding up the resultant values using the Abate-Whitt formula to

Obtaining the CDF

When inverting a LST corresponding to a probability distribution, the Abate-Whitt formula obtains the PDF of the distribution, $f(t)$, by default. We often would like to obtain the CDF, $F(t)$, instead.

To obtain the CDF, we use the derivative property of LSTs:

$$\mathcal{L}^*\{f(t)\}(s) = \mathcal{L}^*\{F'(t)\}(s) = s\mathcal{L}^*\{F(t)\}(s) - F(0^+),$$

where $\mathcal{L}^*\{f(t)\}(s)$ denotes the LST of $f(t)$.

If $F(0^+) = 0$ (which is always the case for the distributions we shall consider), then $\mathcal{L}^*\{F(t)\}(s) = (1/s)\mathcal{L}^*\{f(t)\}(s)$.

Thus, in our LST inversion routine, we simply divide by s to obtain the CDF in place of the PDF. This is implemented at (†) in Algorithm B.2.

obtain the inverse transform. We split these steps into two functions: SETUPILT constructs the set of nodes and weights; INVERTLT evaluates the LST and applies the Abate-Whitt formula.

The LST inversion algorithm can be summarised as follows:

- Set up the inversion (SETUPILT):
  - Specify a set of t values at which we intend to evaluate the inverse LST.
  - Get nodes and weights for each value of t. For the Euler method, these are computed; for the CME method, these are extracted from a file.
- Perform inversion (INVERTLT):
  - Evaluate the LST at all nodes.
  - Apply the Abate-Whitt formula to obtain the inverse LST.

A flow chart of the algorithm is shown in Figure 4.1. Pseudocode and explanations of SETUPILT and INVERTLT can be found in Section B.1.

4.1.2 Motivation

The TTM is a relatively complex LST, due to being constructed from many other components from the fluid model framework. As a result, we shall encounter several hurdles in the process of inverting the LST of the TTM. The biggest

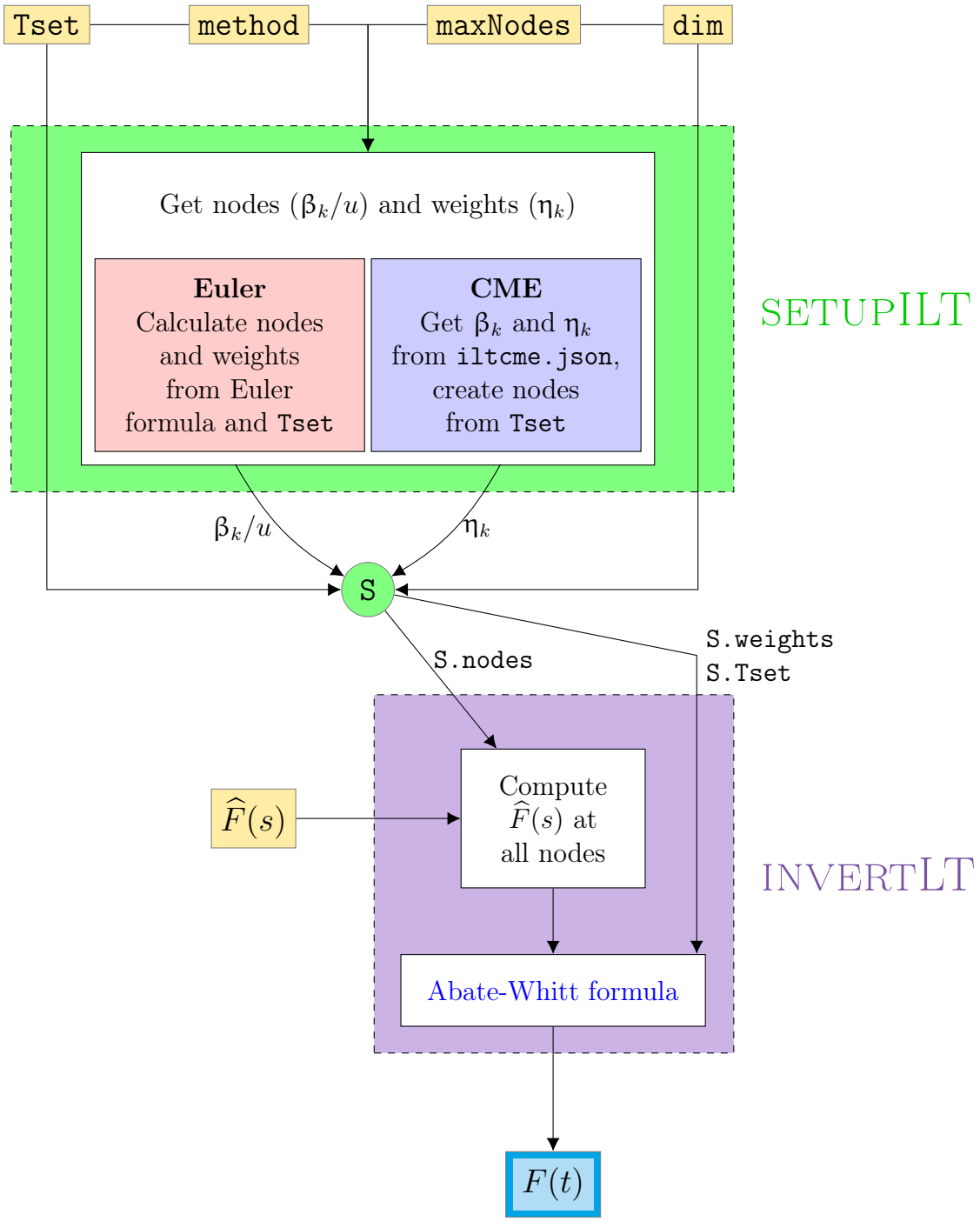


Figure 4.1: Flow chart representation of the procedure for inverting LSTs. SETUPILT sets up the inversion using inputs `Tset`, `method`, `maxNodes` and `dim`, and stores the necessary inversion parameters in `S`. INVERTLT uses `S` to perform the inversion on the LST $\widehat{F}(s)$. The output is $F(t)$, the inverse LST.

problem is that some components which make up the TTM have *point masses*[1] in their associated distributions. As we shall see, point masses severely complicate the inversion process, particularly when the Euler method is used for inversion.

4.2 Inversion hurdles

In this section, we investigate the effect of point masses on the LST inversion process. To do so, we first consider simple LSTs; as we shall see, the same principles apply to fluid model quantities.

Throughout, we use SETUPILT and INVERTLT for all LST inversion.

4.2.1 Point masses

The simplest distribution with a point mass is the distribution of a deterministic random variable with value a. The CDF of such a random variable is the shifted Heaviside function, $H_a : [0, \infty) \to [0, 1]$, defined here as

$$H_a(t) = \begin{cases} 0 & t < a, \\ 1 & t \geq a, \end{cases}$$

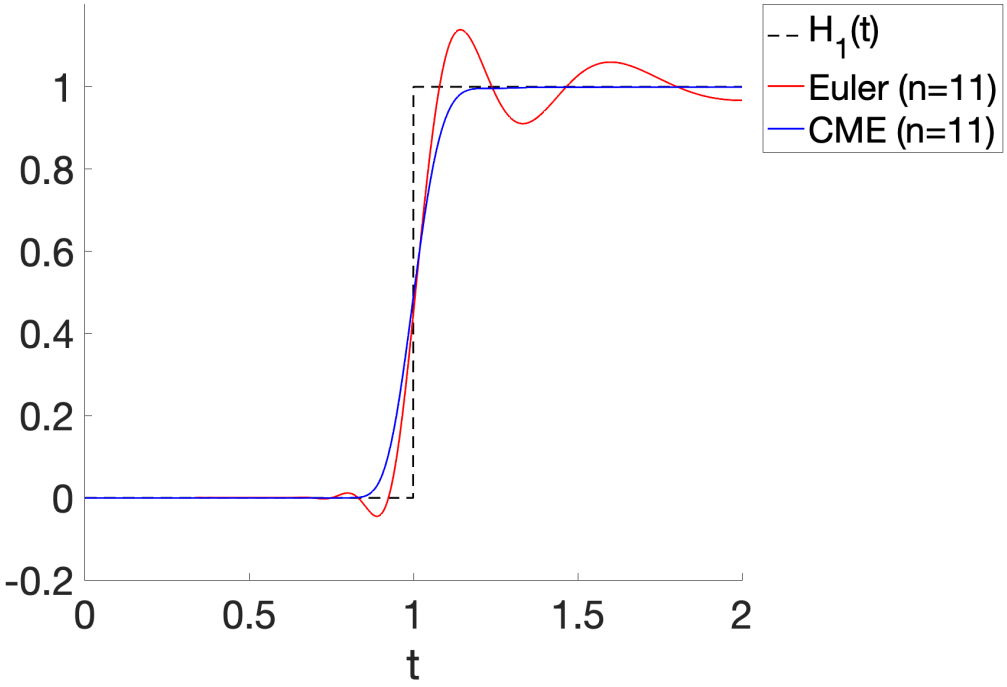
for some $a \in \mathbb{R}$. The LST of $H_a(t)$ is $\widehat{H}_a(s) = e^{-sa}$.

Figure 4.2a shows the results of inverting $\widehat{H}_1(s)$ to obtain $H_1(t)$. Both methods struggle to capture the discontinuity at $t = 1$, but the Euler method is particularly problematic since it goes below 0 and above 1 and oscillates. This is worrying because we are interested in obtaining CDFs with this inversion method, which must be increasing, and lie in the range $[0, 1]$ for all t.

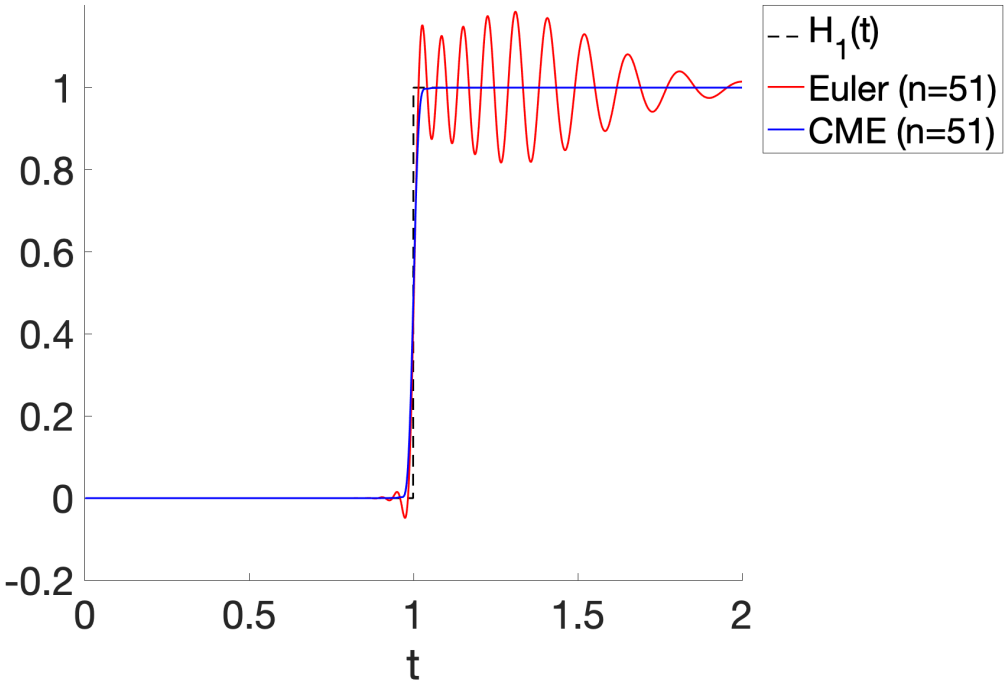
As briefly mentioned in Section 2.3, a major part of Laplace transform inversion relies on approximating the Dirac delta functional; the accuracy of this approximation determines the accuracy of the inversion result. It is clear that the Euler method is affected to a greater degree than the CME method, but both are affected nonetheless. A possible solution is to increase the number of nodes, n, since this uses a higher order approximation to the Dirac delta functional in the Abate-Whitt framework.

Figure 4.2b show the results of the same inversion with $n = 51$ nodes, and we observe that although the CME approximation has greatly improved, the Euler

[1] In a continuous probability distribution, a *point mass* is a finite, positive probability associated with a point a. This is seen as a discontinuity in the associated CDF, and is the reason we consider Laplace-*Stieltjes* transforms rather than Laplace transforms.



(a) $n = 11$ nodes.



(b) $n = 51$ nodes.

Figure 4.2: Approximations to $H_1(t)$ obtained from LST inversion, using the Euler and CME methods.

approximation[2] has significantly worsened after $t = 1$, and is still negative immediately before $t = 1$. A more sophisticated approach may be more helpful.

Removing point masses

To deal with point masses, we make use of the following lemma.

Lemma 4.1. *Let $F(t)$ be a function with a point mass p at $t = a$:*

$$F(t) = \begin{cases} G(t) & 0 \le t \le a, \\ G(t) + p & t > a, \end{cases}$$

for a continuous and increasing function $G(t)$, with $\widehat{F}(s) = \mathcal{L}^*\{F(t)\}(s)$. If $\widehat{G}(s) = \mathcal{L}^*\{G(t)\}(s)$, then $\widehat{F}(s) = pe^{-sa} + \widehat{G}(s)$.

We would prefer to invert $\widehat{G}(s)$, as this LST is likely to cause fewer inversion problems. Thus, to deal with a point mass, we can manipulate the LST of our function during inversion, as follows:

- Evaluate $\widehat{G}(s) = \widehat{F}(s) - pe^{-sa}$ in place of $\widehat{F}(s)$.
- Apply the Abate-Whitt formula to obtain $G(t)$
- Recover $F(t)$ by adding back the point mass: $F(t) = G(t) + p\mathbb{I}(t \ge a)$.

Figure 4.3 shows the result of removing the point mass of $H_1(t)$ at $t = 1$ during the inversion process. Both the Euler and CME methods perform much better. In fact, they both recover the function *exactly*—fair warning, this is unlikely to happen with more complicated functions.

This is an excellent tool for dealing with point masses; however, a large disadvantage is that we need to know precisely the value of the point mass, p, and the time at which it occurs, a. It is also unfit to handle more complicated LSTs, as we shall see shortly.

4.2.2 Dispersed point masses

A key concept that we explore in fluid models is the concatenation of simple paths to form more complicated paths (for an example, see Concatenating paths). In particular, we are interested in adding the random times taken to complete simple paths to find the random time taken to complete a more complicated path. For this reason, we rely heavily on the convolution theorem (Theorem 3.1).

[2]The oscillations that we observe in Figure 4.2a and Figure 4.2b are also observed in the Euler approximation to the Dirac delta functional.

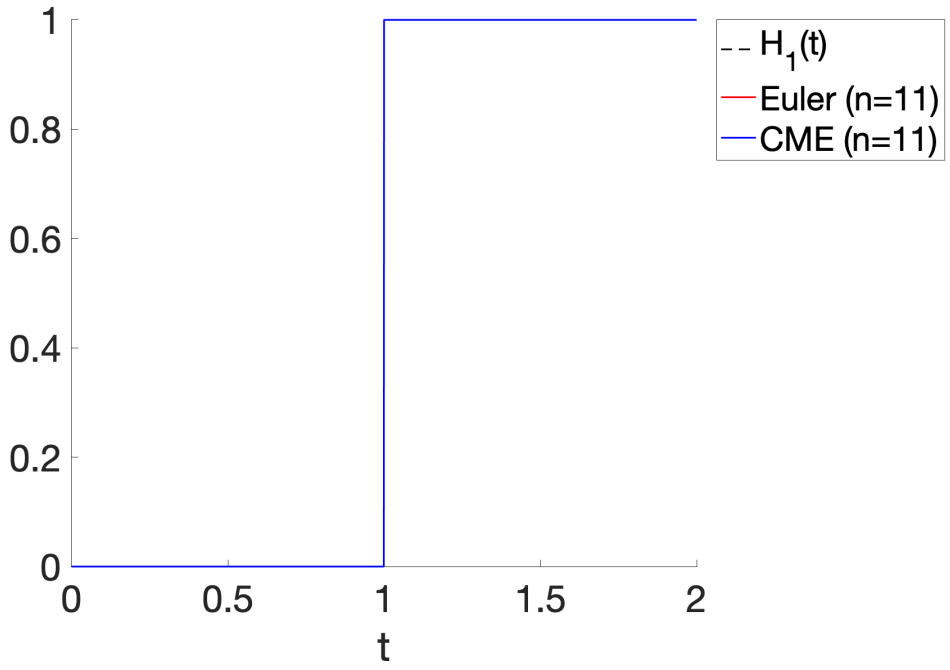


Figure 4.3: Here, the point mass at $t = 1$ is removed, then added back after inversion. Both the Euler and CME methods recover the function exactly.

Consider the case where $X \sim \exp(\lambda)$ and $Y = a$. Then the respective LSTs of X and Y are $\widehat{F}_X(s) = \lambda/(\lambda + s)$ and $\widehat{F}_Y(s) = e^{-sa}$.

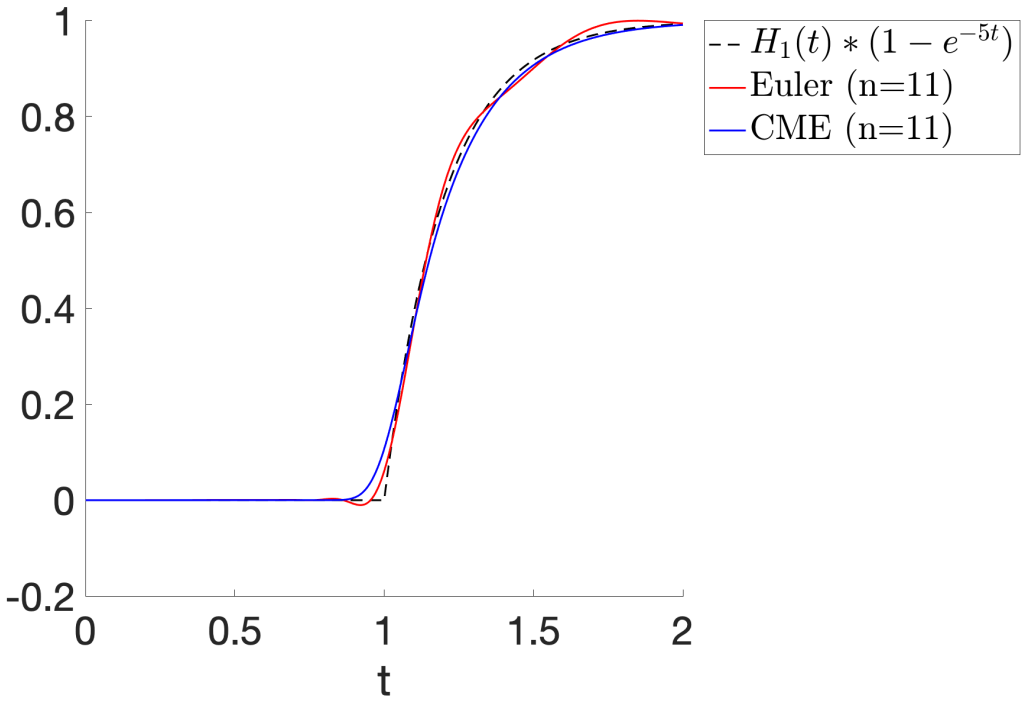
By Theorem 3.1, the LST of $X + Y$ is

$$\widehat{F}_{X+Y}(s) = \frac{\lambda e^{-sa}}{\lambda + s}. \tag{4.1}$$

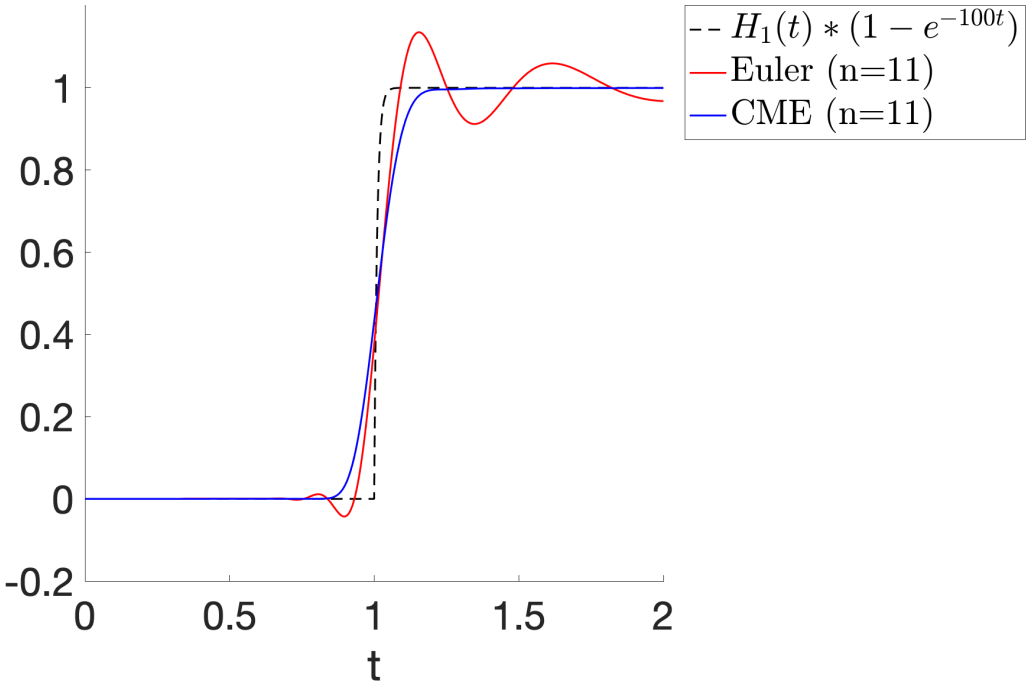
Figure 4.4a shows the result of applying LST inversion to $\widehat{F}_{X+Y}(s)$ when $a = 1$ and $\lambda = 5$. Both methods manage to roughly capture the distribution, although the Euler method still produces slightly negative results immediately before $t = 1$. Figure 4.4b shows the results for $\lambda = 100$, which—particularly for the Euler method—is far less impressive. The result looks very similar to Figure 4.2a, but with one key difference: the random variable $X + Y$ has no point mass.

Despite the fact that $H_1(t)$ has a point mass, the point mass is *dispersed* by the exponential component (this is true for any value of λ). This presents a problem, since the Euler method clearly cannot cope with the sudden jump at $t = 1$, and we cannot use the point mass removal technique from Section 4.2.1 to fix the issue.

Moreover, we are very likely—in hindsight, one could even say *almost sure*—to encounter similar problems in the context of fluid models, and hence, we look to other methods of dealing with such hurdles.



(a) $\lambda = 5$.



(b) $\lambda = 100$.

Figure 4.4: Approximations to $F(t) = H_1(t) * (1 - e^{-\lambda t})$, for $\lambda = 5, 100$, obtained by inverting the LST $\widehat{F}_{X+Y}(s)$ from Eq. (4.1), using the Euler and CME methods ($n = 11$ nodes). These examples illustrate the effect of *dispersed point masses* on LST inversion.

4.2.3 Time-shifting

We now introduce a technique that can be incorporated into the LST inversion routine to deal with sharp changes (discontinuous derivatives of some order) in the CDF that are not point masses, such as dispersed point masses. This technique is known as *time-shifting*, and concerns functions with a sharp change at $t = a$.

In the context of fluid models, we primarily focus on CDFs. Moreover, these CDFs have the same form as in Figure 4.4, where there is some *delay* associated with the CDF. That is, the CDF is zero up until some value a, at which there is a sharp change. We refer to such functions as *delayed CDFs*, or say that a CDF is *delayed by* a. In our case, it is implicit that $a \geq 0$.

Suppose we would like to invert the LST of a delayed CDF with time-shifting. The process of time-shifting can be summarised in three steps:

1. Remove the segment before a (first shift).
2. Invert the LST.
3. Add back the segment before a (second shift).

Step 2 is taken care of, and Step 3 is straightforward since we know the CDF is zero before a, but Step 1 presents a problem. Removing a segment of a function is easy enough, but we must do so by manipulating the LST *only*.

Luckily, a property of LSTs allows us to do this. This property—sometimes called the *second shifting property*—is stated in the following theorem.

Theorem 4.2. *Let $H(t)$ be a function with $\mathcal{L}^*\{H(t)\}(s) = \widehat{H}(s)$, and let*

$$G(t) = \begin{cases} H(t-a) & t > a, \\ 0 & t < a, \end{cases}$$

for some $a \in \mathbb{R}$, with $\mathcal{L}^*\{G(t)\}(s) = \widehat{G}(s)$. Then

$$\widehat{G}(s) = e^{-sa}\widehat{H}(s).$$

In our case, $G(t)$ is the delayed CDF; $\widehat{H}(s)$ is the non-delayed LST that we would like to obtain in Step 1. Immediately, we can see that $\widehat{H}(s) = e^{sa}\widehat{G}(s)$, and hence, we can perform Step 1 of time-shifting.

The entire process is summarised in Figure 4.5. By applying time-shifting when inverting $\widehat{F}(s)$ from Eq. (4.1) (using $n = 11$ nodes), we can obtain much better results, as shown in Figure 4.6 where the telltale oscillations from the Euler method are removed.

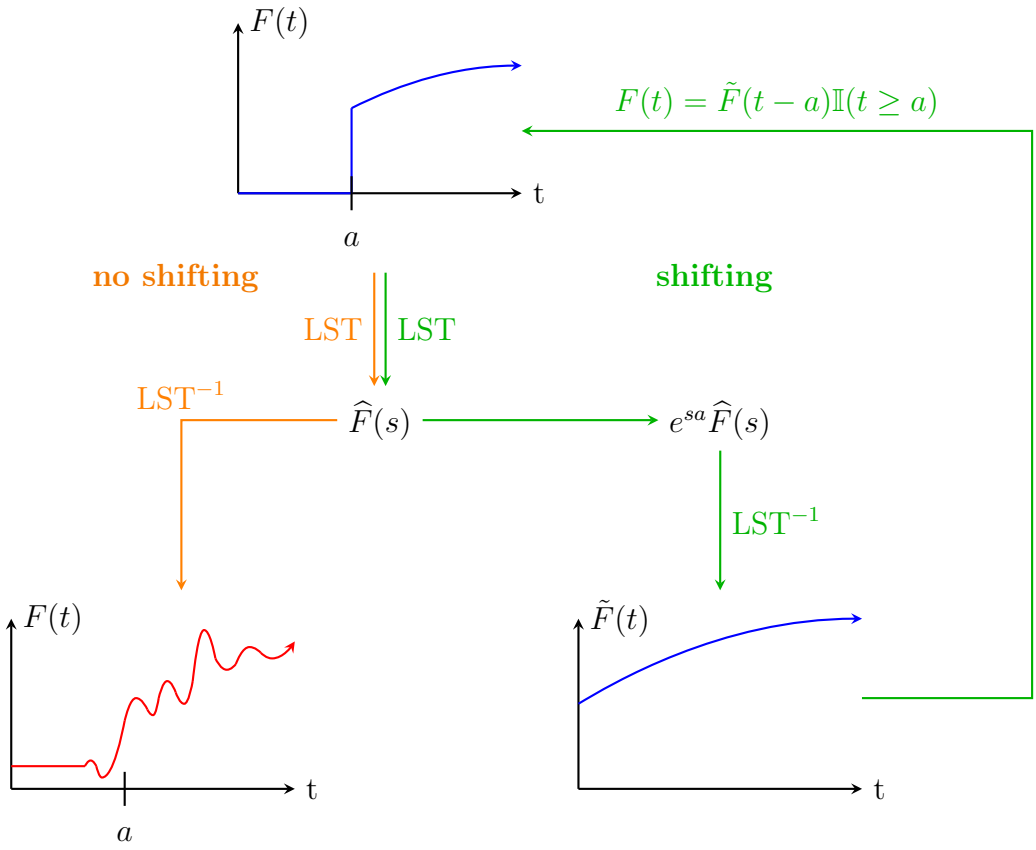
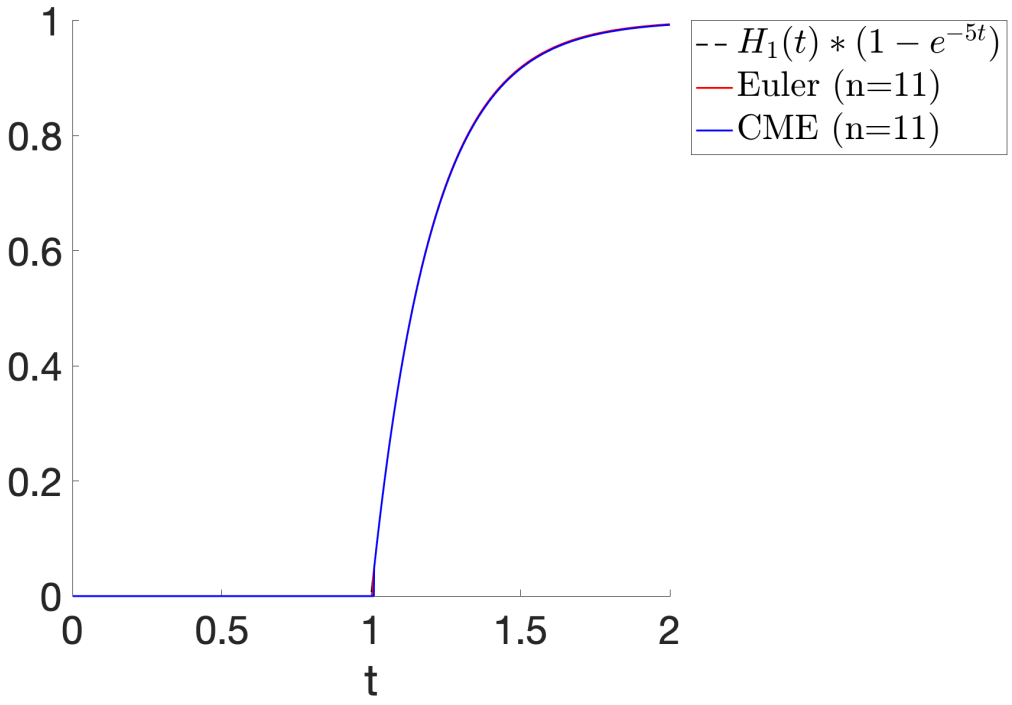
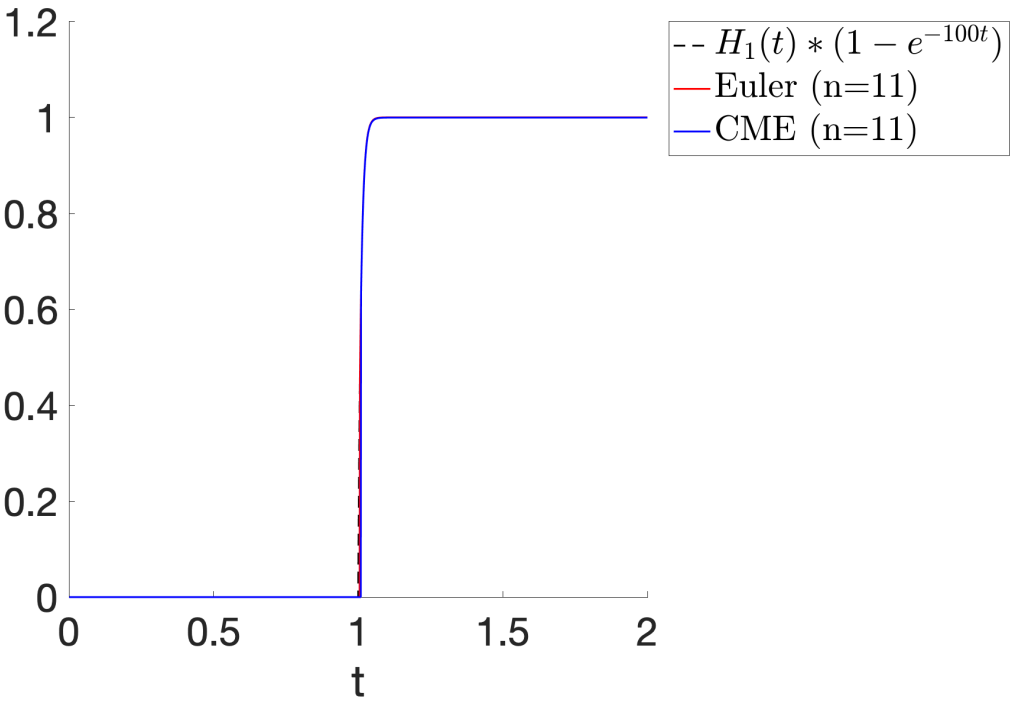


Figure 4.5: Illustration of the time-shifting process. For a CDF delayed by a, using the Euler* method without shifting to invert its LST $\widehat{F}(s)$ yields a bad approximation. We can greatly improve the approximation by using time-shifting. To time-shift, we multiply $\widehat{F}(s)$ by e^{sa}, to remove the segment of $F(t)$ before the discontinuity at a; then, we invert $e^{sa}\widehat{F}(s)$ to obtain a backward-shifted version of $F(t)$, denoted $\tilde{F}(t)$; finally, we shift $\tilde{F}(t)$ forward to recover $F(t)$, via the formula $F(t) = \tilde{F}(t-a)\mathbb{I}(t \geq a)$. *We can use time-shifting with *any* method of LST inversion, but it is most effective when used with the Euler method, which can often produce inverse LSTs resembling the bottom left curve.



(a) $\lambda = 5$.



(b) $\lambda = 100$.

Figure 4.6: Approximations to $F(t) = H_1(t) * (1 - e^{-\lambda t})$, for $\lambda = 5, 100$, obtained by inverting the LST $\widehat{F}_{X+Y}(s)$ from Eq. (4.1), using the Euler and CME methods ($n = 11$ nodes) with time-shifting at $a = 1$. Compared to Figure 4.4 (no time-shifting) the results have greatly improved.

4.3 Inverting the TTM

As we shall see in this section, the TTM is composed of components which have dispersed point masses associated with them. As a result, time-shifting is helpful if we are to use the Euler method for inversion. The CME method is already fairly effective at inverting the TTM without shifting, but it may still be worth time-shifting with the CME method, which is a topic of investigation in Section 4.4.

4.3.1 Point masses in fluid models

The main components in the TTM are the LSTs ${}_z^b\widetilde{\mathbf{\Psi}}(s)$, ${}_0^z\widetilde{\mathbf{\Xi}}(s)$ and ${}_x^b\widetilde{\mathbf{G}}_{--}(s;z)$. The main sub-components of each of these LSTs are ${}_x^y\widehat{\mathbf{G}}(s;z)$ and ${}_x^y\widehat{\mathbf{H}}(s;z)$, which are themselves constructed from $\widehat{\mathbf{G}}(s;z)$ and $\widehat{\mathbf{H}}(s;z)$.

Consider the quantity $\widehat{\mathbf{H}}_{++}(s;1)$: this is the LST of the amount of time taken to fill the fluid buffer from level 0 to level 1. In the canonical model from Chapter 3, this is a 2×2 matrix, since there are 2 elements of $\mathcal{S}_+$, corresponding to genus C and genus D. The (i,j) entry of $\widehat{\mathbf{H}}_{++}(s;1)$ corresponds to starting in genus i and ending in genus j, for $i,j \in \{C,D\}$. Let $c_C^+, c_D^+ > 0$ be the respective fluid input rates corresponding to genus C and genus D during growth.

If $i = D$, then the process can reach level 1 in either genus C or genus D. In the former case, the process must switch into genus C at a random time. In the latter case, the process either remains in genus D for the entire journey, or switches phase multiple times (this happens with very low probability).

If genus D remains dominant for the entirety of the journey, there is a point mass at $t = 1/c_D^+$. There is a positive probability of $e^{-\alpha/c_D^+}$ that genus D remains dominant for the entire journey, and so the journey takes exactly this long.

Similarly, if $i = C_+$ and $j = C_+$, there is a point mass at $t = 1/c_C^+$, corresponding to a path in which the process stays in genus C for the entire journey.

A similar story can be told about $\widehat{\mathbf{G}}_{--}(s;1)$: there are point masses associated with the diagonal elements, corresponding to staying in the same phase for the entirety of the journey from level 1 to level 0.

The off-diagonal elements of $\widehat{\mathbf{H}}_{++}(s;1)$ and $\widehat{\mathbf{G}}_{--}(s;1)$ do not have point masses, since they involve a change of phase which occurs at a random time, and hence the entire journey takes a random length of time.

Quantities such as $\widehat{\mathbf{\Psi}}(s)$, $\widehat{\mathbf{\Xi}}(s)$, $\mathbf{\Lambda}(s)$ and $\mathbf{V}(s)$ do not have point masses. For example, $\widehat{\mathbf{\Psi}}(s)$ is the LST of the time taken to return to an initial level z, starting in an upward phase. It is possible for the process to return at *any* positive time. For any $\varepsilon > 0$, it is possible—albeit very unlikely, if ε is very small—for the process to change into a downward phase and return to z before $t = \varepsilon$.

Recall from Theorem 3.1 that taking sums of independent random variables amounts to taking products of LSTs. In fluid models this is very important, since we are taking a sum of independent random variables when we concatenate conditionally independent paths together (see Matrix convolution example for an example using matrices). For example, consider a path in the unbounded model which begins at level z in a downwards phase, crosses level 0 (down-down path), returns to level 0 from below (down-up path), and then returns to level z (up-up path). The time taken to complete this path is $X + Y + Z$, where:

- X is the time taken to complete a path from level z to level 0. The LST of X is $\widehat{\mathbf{G}}_{--}(s; z)$.
- Y is the time spent below level 0. The LST of Y is $\widehat{\boldsymbol{\Xi}}(s)$.
- Z is the time taken to complete a path from level 0 to level z. The LST of Z is $\widehat{\mathbf{H}}_{++}(s; z)$.

By Theorem 3.1, the LST of $X + Y + Z$ is $\widehat{\mathbf{G}}_{--}(s; z)\,\widehat{\boldsymbol{\Xi}}(s)\,\widehat{\mathbf{H}}_{++}(s; z)$.

The interesting part about this combined LST is that $\widehat{\mathbf{G}}_{--}(s; z)$ and $\widehat{\mathbf{H}}_{++}(s; z)$ have point masses associated with them, but $\widehat{\boldsymbol{\Xi}}(s)$ does not. Therefore, the product will have a dispersed point mass, similar to what we saw in Section 4.2.2.

This motivates the use of time-shifting in the inversion of the TTM: the TTM is constructed from components which contain point masses, such as $\widehat{\mathbf{G}}(s; z)$ and $\widehat{\mathbf{H}}(s; z)$, as well as quantities which disperse these point masses, such as $\boldsymbol{\Lambda}(s)$, $\mathbf{V}(s)$, and the off-diagonal elements of $\widehat{\mathbf{G}}(s; z)$ and $\widehat{\mathbf{H}}(s; z)$. As we saw in Section 4.2.3, time-shifting is an effective method for dealing with dispersed point masses.

There is, however, another problem which arises when we consider shifting the TTM. We know how to shift a function with a dispersed point mass at $t = a$, but what happens when we add or multiply LSTs with such point masses? For example, recall that ${}^{z}_{0}\widetilde{\boldsymbol{\Xi}}(s) = {}^{z}_{0}\widehat{\mathbf{H}}_{-+}(s; z) + {}^{z}_{0}\widehat{\mathbf{G}}_{--}(s; z)\,{}^{z}_{0}\widetilde{\mathbf{W}}(s)\,{}^{z}_{0}\widehat{\mathbf{H}}_{++}(s; 0)$, which contains both sums and products of LSTs. To use time-shifting with such quantities, we require a more sophisticated approach.

4.3.2 Automatic shifting

Let X and Y be random variables with CDFs $F_X(t)$ and $F_Y(t)$ respectively. Let

$$\widehat{F}_X(s) = \mathcal{L}^*\{F_X(t)\}(s),$$
$$\widehat{F}_Y(s) = \mathcal{L}^*\{F_Y(t)\}(s),$$

be the LSTs associated with X and Y, respectively.

Also suppose that $F_X(t)$ is delayed by $t = a_X$, and $F_Y(t)$ is delayed by $t = a_Y$. These delays may be associated with point masses, or dispersed point masses, but the key property is that LST inversion is unstable around these points (as in Figure 4.4). For this reason, we refer to such points as *unstable* points from now on. We assume that a_X and a_Y are the only unstable points.

To invert $\widehat{F}_X(s)$ or $\widehat{F}_Y(s)$, we can use the regular time-shifting technique explored in Section 4.2.3, with a shift of a_X or a_Y, respectively; however, what if we wanted to invert $\widehat{F}_X(s)\widehat{F}_Y(s)$ or $\widehat{F}_X(s) + \widehat{F}_Y(s)$?

Shifting products of LSTs

Recall that $\widehat{F}_X(s)\widehat{F}_Y(s)$ is the LST of $F_X(t) * F_Y(t)$, where $*$ denotes a convolution. Thus, the product $\widehat{F}_X(s)\widehat{F}_Y(s)$ is the LST of $X + Y$, and the corresponding distribution has an unstable point at $a_X + a_Y$, since X and Y are independent. Thus, to invert $\widehat{F}_X(s)\widehat{F}_Y(s)$, we can use time-shifting with a shift of $a_X + a_Y$.

The same methodology holds for a product with any number of multiplicands.

Shifting sums of LSTs

By the linearity of the LST, $\widehat{F}_X(s)+\widehat{F}_Y(s)$ is the LST of $F_X(t)+F_Y(t)$—assuming this is a well-defined CDF, that is, $F_X(\infty) + F_Y(\infty) \leq 1$, which is always true in the case of fluid models.

If $a_X = a_Y$, then $F_X(t) + F_Y(t)$ has a delay of $a_X = a_Y$, and so we can invert $\widehat{F}_X(s) + \widehat{F}_Y(s)$ with a shift of $a_X = a_Y$.

If $a_X \neq a_Y$, then $F_X(t) + F_Y(t)$ has a delay of $\min\{a_X, a_Y\}$, but may also have a sharp change at $\max\{a_X, a_Y\}$. Thus, to use time-shifting here, we must be more careful. To correctly time-shift, we must:

1. Invert $\widehat{F}_X(s)$ with a shift of a_X to obtain $F_X(t)$.
2. Invert $\widehat{F}_Y(s)$ with a shift of a_Y to obtain $F_Y(t)$.
3. Add the results to obtain $F_X(t) + F_Y(t)$.

Shifting products of sums of products of ...

Suppose we are presented with a more complicated expression. For example, let $\widehat{A}(s)$, $\widehat{B}(s)$, $\widehat{C}(s)$, $\widehat{D}(s)$ and $\widehat{E}(s)$ be LSTs with unstable points at $\chi_a, \chi_b, \chi_c, \chi_d$ and χ_e, respectively. We would like to invert the LST

$$\widehat{A}(s)\widehat{B}(s) + \widehat{C}(s)\widehat{D}(s) + \widehat{E}(s). \tag{4.2}$$

We can do this using our previously defined rules, as follows:

- Invert $\widehat{A}(s)\widehat{B}(s)$ with a shift of $\chi_a + \chi_b$, to obtain $x(t)$.
- Invert $\widehat{C}(s)\widehat{D}(s)$ with a shift of $\chi_c + \chi_d$ to obtain $y(t)$.
- Invert $\widehat{E}(s)$ with a shift of χ_e to obtain $z(t)$.
- The final result is $x(t) + y(t) + z(t)$.

This process can be simplified even more if the shifts of two or more summands are equal. For instance, if $\chi_a+\chi_b = \chi_c+\chi_d$, then we can invert $\widehat{A}(s)\widehat{B}(s)+\widehat{C}(s)\widehat{D}(s)$ with a shift of $\chi_a + \chi_b$, rather than in two steps.

In general, we would like to perform the least number of inversions possible, as LST inversion is computationally expensive. In addition, the process gets more complicated if we do more additions and multiplications, for instance, the LST $(\widehat{A}(s) + \widehat{B}(s)\widehat{C}(s))^2(\widehat{D}(s) + \widehat{E}(s))$ adds even more steps to the process.

To make things easier on ourselves, we can semi-automate the process of shifting, meaning that we can invert any such expression where the simplest parts of the expression (i.e., $\widehat{A}(s)$, $\widehat{B}(s)$, etc.) have a single, known, unstable point. The full details of this procedure are outlined in the insert Atoms and molecules I, and an example is given in the insert Atoms and molecules II on the next few pages. We use this shifting procedure, referred to as *atom-based shifting*, for all subsequent time-shifting.

Note that the performance of atom-based shifting is highly dependent on its implementation, and the type of LSTs being inverted. For more information about this, see the insert Atoms and molecules III.

4.3.3 Components of the TTM

Recall from Chapter 3 that the LST of the TTM, starting from phase $i \in \mathcal{S}_-$, is

$$\widehat{F}_{\text{TTM}}(s) = \boldsymbol{e}_i^{\intercal}\, {}_{\zeta}^{b}\widetilde{\mathbf{G}}_{--}(s;b)\left(\mathbf{I} - \int_0^{\tau} e^{-su}\boldsymbol{\xi}(u)\,du\, {}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right)^{-1} e^{-s\tau}\left(\mathbf{1}_- - \int_0^{\tau}\boldsymbol{\xi}(u)\,du\,\mathbf{1}_+\right),$$

where $\boldsymbol{\xi}(t)$ is the inverse LST of ${}_{0}^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)$.

As a result, the main components of the TTM are ${}_{\zeta}^{b}\widetilde{\mathbf{G}}_{--}(s;b)$, ${}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)$ and $\boldsymbol{\xi}(t)$ (obtained via ${}_{0}^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)$). We begin by using atom-based shifting to obtain $\boldsymbol{\xi}(t)$: this is a critical component of the TTM, since it governs the geometrically distributed number of re-visits to level ζ, as we shall see shortly.

Atoms and molecules I: methodology

When dealing with fluid models, we often encounter LSTs which are complicated sums and products of other LSTs, such as in Eq. (4.2). Importantly, the sub-components are often relatively simple to invert; it is the combination of such LSTs which is difficult. This is what *atom-based shifting* aims to solve.

Atom-based shifting is introduced here in terms of scalar LSTs, but can be extended to matrix LSTs for use with fluid models (although inverting matrix LSTs can introduce complications, as discussed in Shifting disclaimer).

Atom-based shifting is an object-oriented implementation of the theory from Section 4.3.2, introducing two kinds of objects: *atoms* and *molecules*. Atoms are 'simple' LSTs that we can invert with time-shifting. Molecules represent sums of such LSTs. They are defined as follows:

- An *atom* is an object `A` with an associated function `A.fun`, and shift `A.sh`. The function `A.fun` is a LST $\widehat{A}(s)$; the shift `A.sh` is a non-negative number $a \geq 0$ (if $a = 0$, we do not need to shift during inversion).
  - We can invert the atom `A` by regular LST inversion, with a shift of a, to obtain the inverse LST $A(t)$.
  - Multiplying atoms produces an atom. If `A` and `B` are atoms, then `A`×`B` is an atom with LST $\widehat{A}(s)\widehat{B}(s)$, and shift $a + b$.
  - For equal shifts, adding atoms produces an atom. If `A` and `B` are atoms with respective functions $\widehat{A}(s)$ and $\widehat{B}(s)$, and both with shift a, then `A`+`B` is an atom with LST $\widehat{A}(s) + \widehat{B}(s)$, and shift a.
  - For unequal shifts, adding atoms produces a *molecule*.
- A *molecule* is an object `M` with a collection of atoms, `M.atoms`. These atoms represent the summands in a sum of LSTs.
  - Adding molecules produces a new molecule. If `M` and `N` are molecules, then `M`+`N` contains the atoms of both `M` and `N`.
  - Multiplying molecules produces a new molecule. If `M` and `N` are molecules, then the atoms of `M`×`N` are calculated from the polynomial expansion of `sum(M.atoms)`×`sum(N.atoms)`.
  - Molecules are simplified if any of their atoms have the same shift.
  - Molecules are inverted by inverting their individual atoms and adding the results.

Atoms and molecules II: example

Let `M` and `N` be molecules with `M.atoms = {A,B}` and `N.atoms = {C,D}`, where the atoms `A,B,C,D` have respective LSTs $\widehat{A}(s), \ldots, \widehat{D}(s)$, and shifts $a, \ldots, d$. Suppose we would like to invert `M×N`. Also suppose that $a+c = b+d$.

The molecule `M×N` corresponds to the LST $(\widehat{A}(s)+\widehat{B}(s))(\widehat{C}(s)+\widehat{D}(s))$. We use polynomial expansion to simplify this expression:

$$\left(\widehat{A}(s)+\widehat{B}(s)\right)\left(\widehat{C}(s)+\widehat{D}(s)\right) = \widehat{A}(s)\widehat{C}(s)+\widehat{A}(s)\widehat{D}(s)+\widehat{B}(s)\widehat{C}(s)+\widehat{B}(s)\widehat{D}(s).$$

This is a new molecule `MN` with four atoms: `AC`, `AD`, `BC` and `BD`, with respective LSTs $\widehat{A}(s)\widehat{C}(s)$, $\widehat{A}(s)\widehat{D}(s)$, $\widehat{B}(s)\widehat{C}(s)$ and $\widehat{B}(s)\widehat{D}(s)$, and shifts $a+c$, $a+d$, $b+c$ and $b+d$.

Since $a+c = b+d$, we can simplify `AC+BD` into an atom with LST $\widehat{A}(s)\widehat{C}(s)+\widehat{B}(s)\widehat{D}(s)$ and shift $a+c$. Subsequently, `MN` is a molecule with three atoms.

To invert `MN`, we invert each of these three atoms (with their respective shifts) to obtain $x(t)$, $y(t)$ and $z(t)$, and add the results to obtain the inverse LST $x(t)+y(t)+z(t)$.

Obtaining $\boldsymbol{\xi}(t)$

Recall from Eq. (2.27) that

$$\overset{\zeta}{_0}\widetilde{\boldsymbol{\Xi}}(s) = \overset{\zeta}{_0}\widehat{\mathbf{H}}_{-+}(s;\zeta) + \overset{\zeta}{_0}\widehat{\mathbf{G}}_{--}(s;\zeta)\,\overset{\zeta}{_0}\widetilde{\mathbf{W}}(s)\,\overset{\zeta}{_0}\widehat{\mathbf{H}}_{++}(s;0). \tag{4.3}$$

Unfortunately, Eq. (4.3) does not express the LST in simple enough terms for atom-based shifting to apply. The goal is to simplify Eq. (4.3) into *atoms*. Let us consider $\overset{\zeta}{_0}\widehat{\mathbf{H}}_{++}(s;\zeta)$ first. Recall from Eq. (2.25) that

$$\overset{y}{_0}\widehat{\mathbf{H}}_{++}(s;z) = \left[\widehat{\mathbf{H}}_{++}(s;y-z) - \widehat{\mathbf{G}}_{+-}(s;z)\widehat{\mathbf{H}}_{-+}(s;y)\right]\sum_{m=0}^{\infty}\left[\widehat{\mathbf{G}}_{+-}(s;y)\widehat{\mathbf{H}}_{-+}(s;y)\right]^m.$$

Here, $y=\zeta$ and $z=0$. Thus, using the fact that $\widehat{\mathbf{H}}_{-+}(s;0) = \widehat{\boldsymbol{\Xi}}(s)$, we have that

$$\overset{\zeta}{_0}\widehat{\mathbf{H}}_{++}(s;0) = \left[\widehat{\mathbf{H}}_{++}(s;\zeta) - \widehat{\mathbf{G}}_{+-}(s;\zeta)\widehat{\boldsymbol{\Xi}}(s)\right]\sum_{m=0}^{\infty}\left[\widehat{\mathbf{G}}_{+-}(s;\zeta)\widehat{\mathbf{H}}_{-+}(s;\zeta)\right]^m. \tag{4.4}$$

This expression contains sums and products of $\widehat{\mathbf{G}}(s;\zeta)$ and $\widehat{\mathbf{H}}(s;\zeta)$ only, and thus we can use atom-based shifting! The only problem is the infinite sum.

Atoms and molecules III: computational issues

The performance of atom-based shifting can be severely hindered if implemented naïvely. In particular, this can happen if the LST associated with an atom is particularly slow to compute, and this atom is used to construct many subsequent atoms or molecules.

For example, consider the atom `P` with LST $\widehat{\mathbf{\Psi}}(s)$ and shift 0. Now consider the atom `Q` defined as $\texttt{P}^{100}$, i.e., $\texttt{P}\times\texttt{P}\times\ldots\times\texttt{P}$. To invert `Q`, a naïve algorithm would evaluate $\widehat{\mathbf{\Psi}}(s)$ 100 times for each value of s. It would be far more efficient to evaluate $\widehat{\mathbf{\Psi}}(s)$ once, then take this value to the 100^{th} power.

This problem arises naturally when inverting the TTM, since many quantities are derived from $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$, which require the solving of Sylvester equations for each value of s, and are slow to compute as a result. Thus, in a naïve implementation, inverting the TTM is almost intractable due to the (unnecessarily) repeated computations of $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$.

A workaround is to pre-compute $\widehat{\mathbf{\Psi}}(s)$ and $\widehat{\mathbf{\Xi}}(s)$ at all necessary values of s, and then pass these values to the corresponding atoms, such that the atoms simply return the values which have already been computed.

Unfortunately, this causes another issue: the necessary s values are derived from the set of t values at which we intend to evaluate the inverse LST; but the t values depend on the amount of time-shifting applied, and we must construct the atom first to work out the required time-shifts—a circular construction!

A secondary workaround is to construct a 'lightweight' atom which does not pre-compute, but contains the necessary shifts. The lightweight atom is then used to construct the set of t values, which in turn constructs the set of s values. We can then use *this* set of s and t values to construct the real atom.

This is a major flaw of atom-based shifting: the primary aim of the method is to reduce the amount of effort required to shift complex LSTs, but in doing so, we are required to take many extra steps to make computations tractable!

The infinite sum

There are multiple instances where we are faced with matrix inverses of the form $(\mathbf{I}-\mathbf{X})^{-1}$, which equate to the infinite sum $\sum_{m=0}^{\infty}\mathbf{X}^m$. If we were to authentically apply atom-based shifting, we would need to consider every term of this sum, since each term has a *different shift*.

Referring back to our previous example, $\widehat{\mathbf{G}}_{+-}(s;\zeta)$ and $\widehat{\mathbf{H}}_{-+}(s;\zeta)$ have dispersed point masses at ζ/c_D^- and ζ/c_D^+ respectively, and thus the product has a dispersed point mass at $\zeta/c_D^- + \zeta/c_D^+$. Similarly, $\left[\widehat{\mathbf{G}}_{+-}(s;\zeta)\widehat{\mathbf{H}}_{-+}(s;\zeta)\right]^m$ has a dispersed point mass at $m(\zeta/c_D^- + \zeta/c_D^+)$.

To deal with such infinite sums, we could simply truncate the sum at some large value of m, but this would lead to a lot of LST inversions. Instead, we can create a better approximation. To do this, write the matrix inverse as follows:

$$\begin{aligned}\sum_{m=0}^{\infty}\mathbf{X}^m &= \mathbf{I}+\sum_{m=1}^{\infty}\mathbf{X}^m\\ &= \mathbf{I}+\mathbf{X}\sum_{m=1}^{\infty}\mathbf{X}^{m-1}\\ &= \mathbf{I}+\mathbf{X}\sum_{m=0}^{\infty}\mathbf{X}^{m}\\ &= \mathbf{I}+\mathbf{X}\left(\mathbf{I}-\mathbf{X}\right)^{-1}.\end{aligned}$$

Now, instead of expanding the matrix inverse, we apply atom-based inversion to the expression $\mathbf{I}+\mathbf{X}\left(\mathbf{I}-\mathbf{X}\right)^{-1}$, where $\mathbf{I}$ and $\mathbf{X}$ are atoms, and we approximate $(\mathbf{I}-\mathbf{X})^{-1}$ as one. The atom $\mathbf{X}$ uses its regular shift, and the atom $(\mathbf{I}-\mathbf{X})^{-1}$ has a shift of 0 (this is where the approximation comes in).

Essentially, we have rearranged the expression so that the matrix inverse stays intact, but such that we can still apply more complex shifting techniques.

To increase our accuracy, we could go even further:

$$\begin{aligned}\sum_{m=0}^{\infty}\mathbf{X}^m &= \mathbf{I}+\mathbf{X}+\mathbf{X}^2\left(\mathbf{I}-\mathbf{X}\right)^{-1}\\ &\vdots\\ &= \sum_{m=0}^{M}\mathbf{X}^m+\mathbf{X}^{M+1}\left(\mathbf{I}-\mathbf{X}\right)^{-1},\end{aligned}$$

for some $M\geq -1$. When $M=-1$, we simply use the expression $(\mathbf{I}-\mathbf{X})^{-1}$ with a shift of 0.

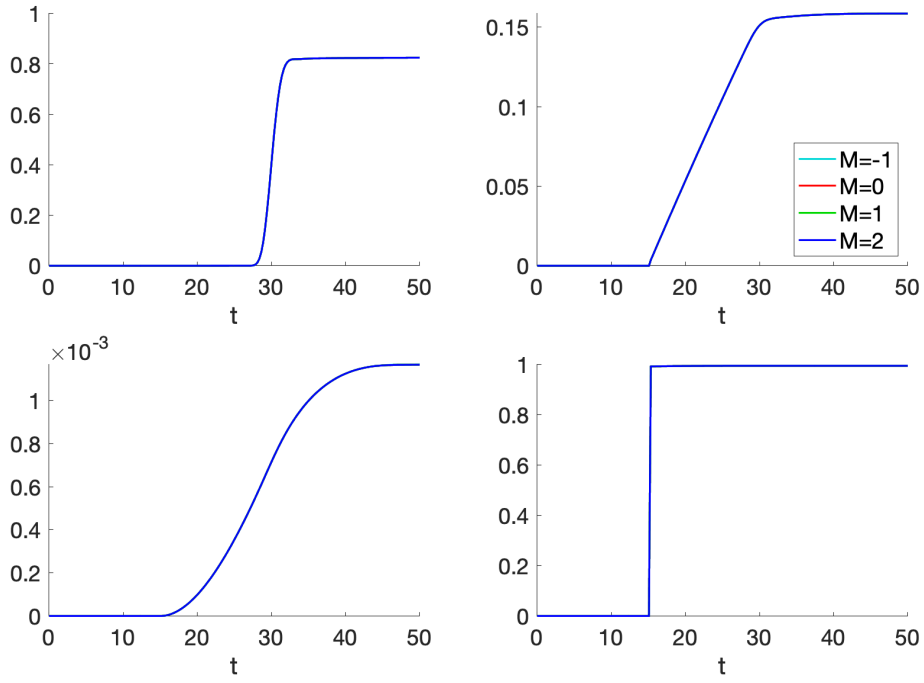


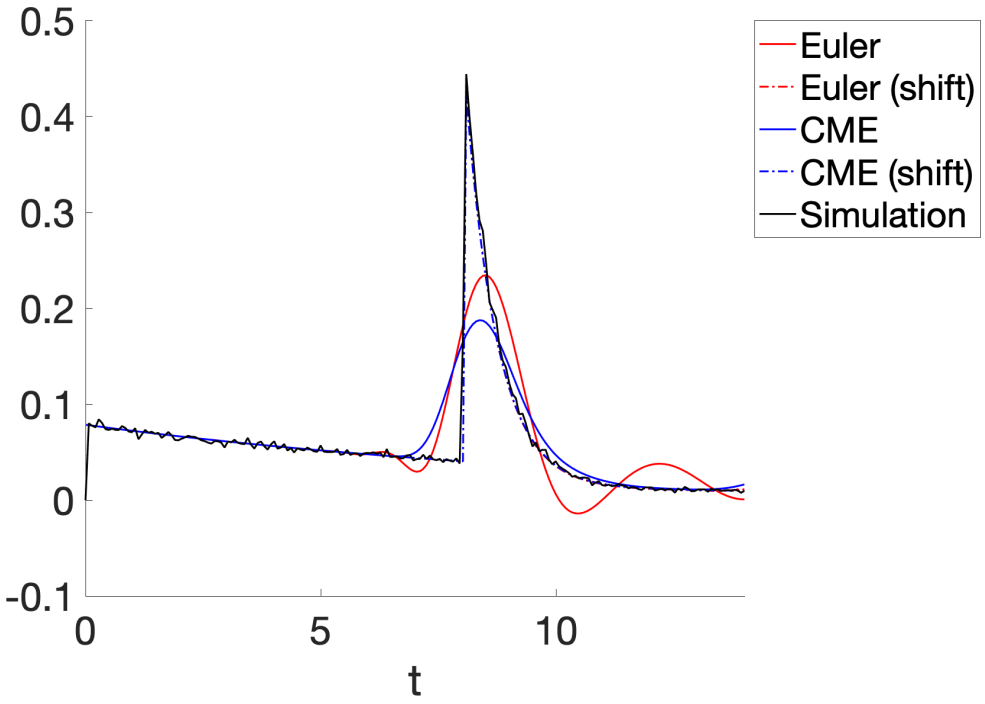
Figure 4.7: CDFs corresponding to the four elements of the 2×2 matrix LST ${}_0^{\zeta}\widehat{\mathbf{H}}_{++}(s;0)$, obtained via LST inversion with atom-based shifting, using the CME method with $n = 11$ nodes. Each individual line represents a different value of M, which is the degree to which we expand the infinite sum in Eq. (4.4). The lines are almost identical, indicating that a value of $M = -1$, corresponding to taking the matrix inverse as an atom with a shift of 0, is sufficient for our purposes.

Figure 4.7 shows the CDFs corresponding to each element of the 2×2 matrix LST ${}_0^{\zeta}\widehat{\mathbf{H}}_{++}(s;0)$, obtained via LST inversion, using the shifted CME method with $n = 11$ nodes and $M \in \{-1, 0, 1, 2\}$. The CDFs are identical, indicating that a value of $M = -1$ is sufficient for inversion. This is the case for many other fluid quantities, and when the Euler method is used (not shown here). Since a value of $M = -1$ requires the least number of LST inversions, we use it from now on.

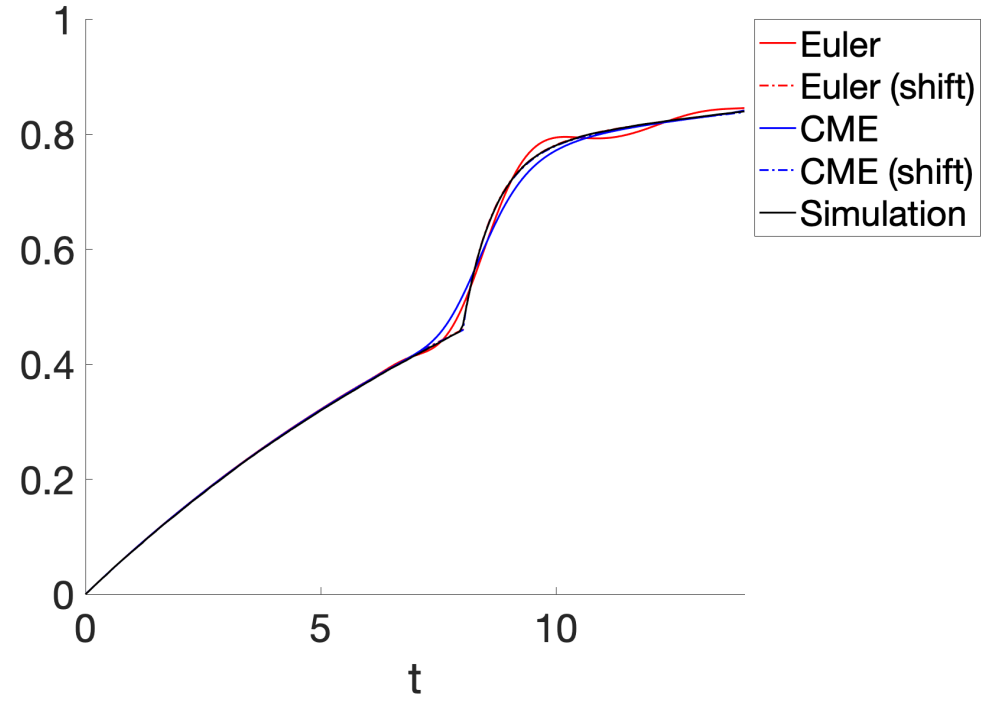
Thus, we can decompose ${}_0^{\zeta}\widehat{\mathbf{H}}_{-+}(s;\zeta)$ into simple LSTs. The same method can be used to decompose ${}_0^{\zeta}\widehat{\mathbf{H}}_{-+}(s;\zeta)$, ${}_0^{\zeta}\widehat{\mathbf{G}}_{--}(s;\zeta)$ and ${}_0^{\zeta}\breve{\mathbf{W}}(s)$, allowing us to decompose ${}_0^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)$ into simple LSTs, and hence, obtain $\boldsymbol{\xi}(t)$ with atom-based shifting.

Observations about $\boldsymbol{\xi}(t)$

In Figure 4.8, we plot the PDF $\boldsymbol{\xi}(t)$ and the corresponding CDF $\int_0^t \boldsymbol{\xi}(u)\,du$, obtained via LST inversion with the Euler and CME methods ($n = 11$ nodes), with and without shifting. Rather than plotting all elements of the matrix, we take the sum of the elements in the first row for each distribution, corresponding to starting in genus C and ending in any genus. For comparison, we plot the



(a) $\boldsymbol{\xi}(t)$ for $t \in [0, \tau]$. Euler (shifted) and CME (shifted) PDFs are almost identical. Both shifted PDFs appear to capture the sharp change at $t = 8$ very well.



(b) $\int_0^t \boldsymbol{\xi}(u)\, du$ for $t \in [0, \tau]$. Euler (shifted) and CME (shifted) CDFs are very close to the simulated CDF.

Figure 4.8: PDF and CDF obtained by inverting ${}_0^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)$ with the Euler and CME methods ($n = 11$ nodes), with and without shifting. We plot the sum of all elements in the first row of each distribution, corresponding to starting in genus C and ending in any genus. Black lines indicate simulated distributions (simulations use Algorithm B.4 with 100,000 samples).

	Simulation	Euler	Euler (shift)	CME	CME (shift)
$\boldsymbol{e}_1^\intercal(\mathbf{1} - \boldsymbol{\Xi}^*\mathbf{1})$	0.1601	0.1478	0.1835	0.1554	0.1816

Table 4.1: Table of $\boldsymbol{e}_1^\intercal(\mathbf{1} - \boldsymbol{\Xi}^*\mathbf{1})$, which is the probability of mortality on a journey below ζ, starting at level ζ in genus C, before the process returns to ζ. Here, $\boldsymbol{e}_1^\intercal$ denotes the (row) unit vector with a 1 in the position corresponding to genus C, and zeros everywhere else. The quantity $\boldsymbol{\Xi}^*$ is the integral of the density shown in Figure 4.8a, approximated using trapezoidal integration.

simulated distributions, obtained with Algorithm B.4 and 100,000 samples. The simulated PDF is the approximate derivative of the simulated CDF; as a result, there is some visible noise in the PDF due to this approximation.

The shifted distributions very closely resemble the simulated ones: we observe in Figure 4.8a that the shifted methods appear to capture the sharp jump at $t = 8$ (corresponding to the minimum time required to return to ζ after hitting level 0) very well; in Figure 4.8b, the shifted methods are almost identical to the simulated CDF. The unshifted inversion methods struggle much more with obtaining the PDF than they do obtaining the CDF. In particular, the unshifted Euler method produces negative values for the PDF.

From both figures, it appears obvious that the shifted Euler and CME methods are superior at inverting $\boldsymbol{\xi}(t)$; however, there is more to be seen here.

Recall from Section 3.3.2 that the final component of the TTM contains the quantity $\boldsymbol{\Xi}^* = \int_0^\tau \boldsymbol{\xi}(u)\,du$. This is the matrix of probabilities that the process returns to ζ, starting from ζ, before mortality occurs. Thus, $\boldsymbol{\Xi}^*\mathbf{1}$ is the probability of returning to ζ before mortality occurs, starting from ζ and ending in any phase. Since we are not interested in what phase mortality occurs in, $\mathbf{1} - \boldsymbol{\Xi}^*\mathbf{1}$ is the probability of mortality occurring before returning to ζ, starting from ζ. Since $c_-^C = c_-^D$ in our case, the initial phase does not have much impact on the mortality process, which is why we assume here that the process starts in genus C. Hence, $\boldsymbol{e}_1^\intercal(\mathbf{1} - \boldsymbol{\Xi}^*\mathbf{1}) = 1 - \boldsymbol{e}_1^\intercal\boldsymbol{\Xi}^*\mathbf{1}$ is the probability of mortality below ζ, starting from genus C and ending in any genus, where $\boldsymbol{e}_1^\intercal$ is the (row) unit vector with a 1 in the position corresponding to genus C.

Table 4.1 tabulates these values for all methods, obtained by integrating the approximations to $\boldsymbol{\xi}(t)$ from Figure 4.8a via trapezoidal integration. The unshifted CME method provides the best approximation. The unshifted Euler method underestimates the probability, but we expect a degree of error due to the oscillations in Figure 4.8a. Importantly, we observe that the shifted Euler and shifted CME methods overestimate the simulated probability of mortality more than the unshifted methods, despite appearing to produce the best estimate of $\boldsymbol{\xi}(t)$.

The underlying reason for the differing values in Table 4.1 is the quantity $\boldsymbol{\Xi}^*$. In Figure 4.8a, we plot the first row sum of $\boldsymbol{\xi}(t)$, which superposes two PDFs, corresponding to $[\boldsymbol{\xi}(t)]_{1,1}$ and $[\boldsymbol{\xi}(t)]_{1,2}$. These are shown individually in Figure 4.9, for $t \in [7.5, 8.5]$. The PDF in Figure 4.9a is not too interesting, but in Figure 4.9b, we see that the shifted PDFs do not capture the sharp change at $t = 8$ as well as we thought. This explains why the shifted methods produce smaller values of $\boldsymbol{\Xi}^*$ than the simulated value, but what is the cause of this inaccuracy?

Recall that to invert a LST, we must evaluate the LST at nodes $s = \beta_k/t$. When inverting a LST with a shift of a, the first shift involves removing the segment before a: in practice, this means if we specify a set $\{t_1, \ldots, t_m\}$ of t values, we actually evaluate the LST at $\{t_p - a, \ldots, t_m - a\}$, where $t_p = \min_i\{t_i > a\}$. We also multiply the LST by the value e^{sa}, completing Step 1 from Section 4.2.3.

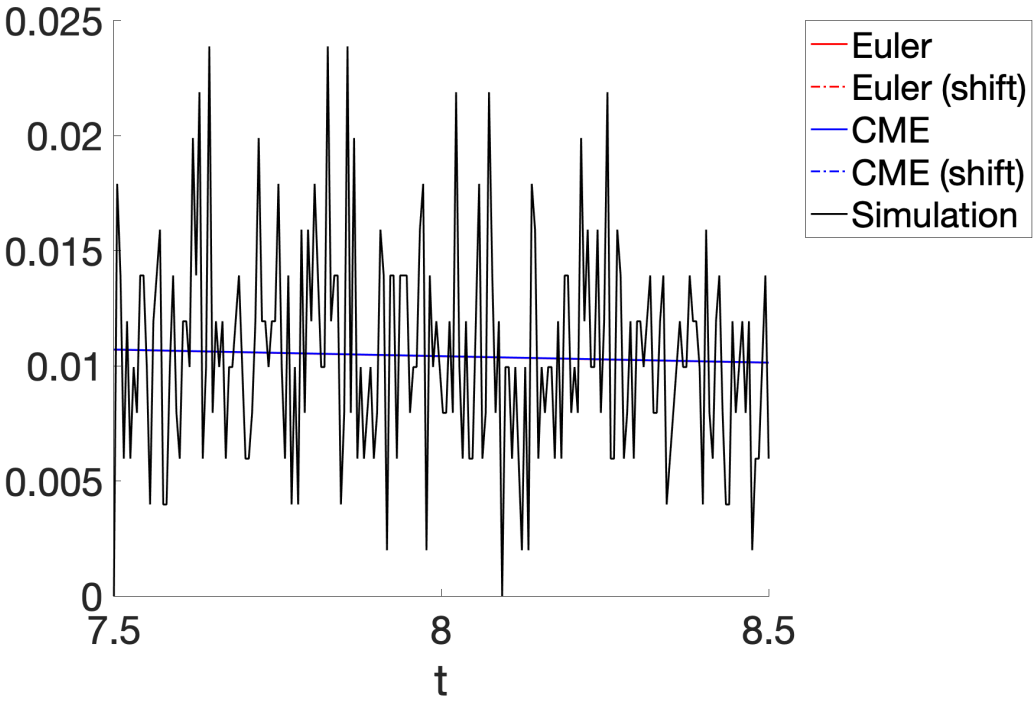
Using the Abate-Whitt formula, we cannot evaluate an LST at $t = 0$, but we can at $t = \varepsilon$, for some small $\varepsilon > 0$. In this case, $s = \beta_k/\varepsilon$, and consequently e^{sa}, is very large. Similarly, e^{-sa}, which is contained in the LST, is very small. In the extreme case, these numbers are so small and large that they are considered to be zero and infinity, respectively. Thus, the product of e^{sa} and the LST is undefined.

It is common to define such products as *NaN* (Not a Number), which is the case in Matlab, but unfortunately, these NaN values completely break the inversion process since we cannot use them in the Abate-Whitt formula. A simple fix is to set the NaN values to zero before applying the Abate-Whitt formula, which is what we do. This makes the inversion process usable, but the inverse LST has slightly inaccurate values to the right of a, once the second shift has been applied. This is the cause of the inaccuracy observed in Figure 4.9b, and why the shifted methods produce smaller values of $\boldsymbol{\Xi}^*$—and hence, larger values of $\boldsymbol{e}_1^\intercal(\mathbf{1} - \boldsymbol{\Xi}^*\mathbf{1})$.

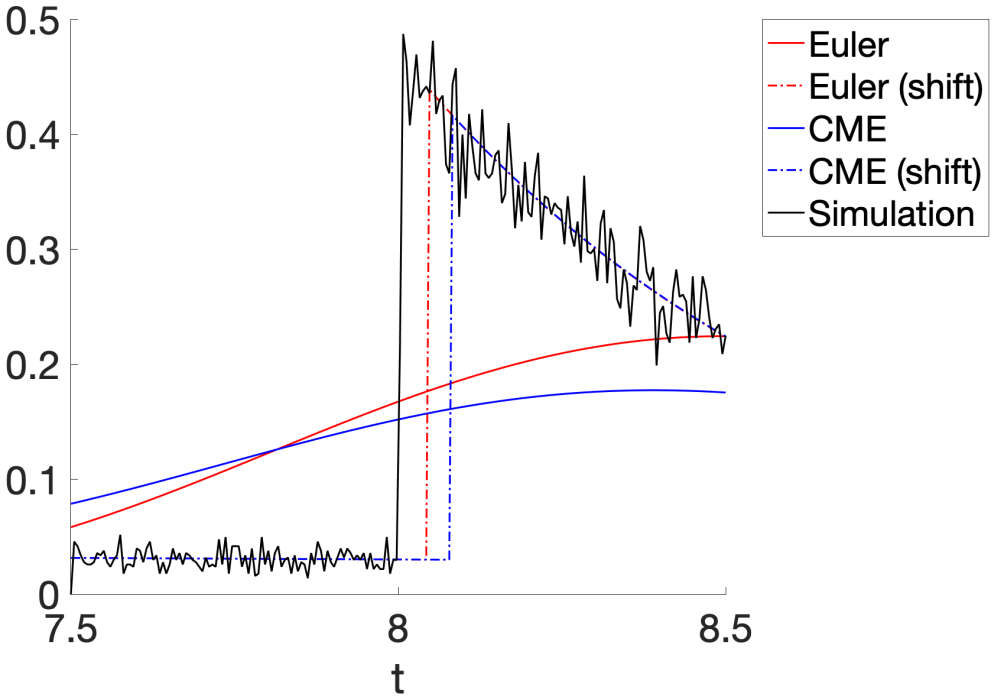
This margin of error may not seem important, but as we shall see, this makes quite a large difference in the inversion of the TTM. Moreover, it is worth noting that for every method, there is some error associated with the trapezoidal integration (for shifting methods, this can accentuate the aforementioned issues). This error can be reduced by increasing the number of points used, although this increases the number of LST inversions performed, slowing down computations.

The rest of the TTM

To obtain ${}_\zeta^b\widetilde{\mathbf{G}}_{--}(s; b)$ and ${}_\zeta^b\widetilde{\boldsymbol{\Psi}}(s)$, we use the exact same method. Each quantity can be decomposed into $\widehat{\mathbf{G}}(s; z)$ and $\widehat{\mathbf{H}}(s; z)$ components, and as a result we can form molecules for each of these quantities. To obtain the TTM, we simply multiply these molecules and invert the result. Finally, we can invert the LST of the TTM.



(a) $[\boldsymbol{\xi}(t)]_{1,1}$, for $t \in [7.5, 8.5]$, corresponding to starting in genus C and ending in genus C.



(b) $[\boldsymbol{\xi}(t)]_{1,2}$, for $t \in [7.5, 8.5]$, corresponding to starting in genus C and ending in genus D. Upon closer inspection, we see that the shifted methods do not quite capture the time of the sharp change at $t = 8$. As a result, the integral of these PDFs, $\boldsymbol{\Xi}^*$, is less than the true value.

Figure 4.9: $(1, 1)$ and $(1, 2)$ elements of $\boldsymbol{\xi}(t)$ with the Euler and CME methods ($n = 11$ nodes), with and without shifting. Black lines indicate simulated distributions (simulations use Algorithm B.4 with 100,000 samples).

Shifting disclaimer

We defined atom-based shifting in terms of scalar LSTs, but when dealing with fluid models, we frequently encounter matrix LSTs. In addition, the elements of these matrices may have different shifts. For instance, in $\widehat{\mathbf{H}}_{++}(s;z)$, the $(1,1)$ element has a different shift than the $(2,2)$ element, since genus C has a smaller growth rate than genus D.

To simplify the process of atom-based shifting, we do not keep track of the individual shifts associated with each component of a matrix LST, but rather associate a single shift—the minimum shift—with each matrix. In the case of $\widehat{\mathbf{H}}_{++}(s;z)$, this is a shift of z/c_D^+, which is the first unstable point of $\widehat{\mathbf{H}}_{++}(s;z)$.

This does have consequences. In particular, after the initial discontinuity, shifting methods do not capture subsequent discontinuities as well.

We could certainly increase the accuracy of the method by keeping track of individual shifts, but this would come at the cost of increased complexity and memory usage, since we are required to specify nm sets of t values for a $n \times m$ matrix LST when using element-wise shifting. Moreover, using the minimum time is often sufficient in practice, as we shall demonstrate.

Final note on shifting

Before moving on to the TTM, we briefly mention an approximation used to simplify the time-shifting process, and cut down on computational time and memory usage. For the rest of the chapter, we take the *minimum* shift required for each matrix quantity, rather than the individual shifts for each component of the matrix. The above insert Shifting disclaimer discusses this in more detail.

4.4 Analysis of the TTM

In this section, we invert the LST of the TTM using the methods covered thus far, and assess the strengths and weaknesses of atom-based shifting. In particular, we begin to focus on the differences between the Euler and CME methods.

We are interested in comparing the performance of the Euler and CME methods. So far, the CME method has proven to be superior in most cases, but using shifting with the Euler method may be able to improve the performance of the method. Moreover, if the CME method does prove to be superior, then is it worth using shifting with the CME method, rather than just increasing the number of nodes, in order to increase accuracy?

Simulated distributions

In this chapter, and in future chapters, we simulate various fluid paths to obtain the distributions of the time taken for these paths to be completed. The underlying algorithm used to do this is Algorithm B.4, which simulates a 'simple' path and returns the time elapsed. We concatenate such paths to obtain more complicated distributions, such as the TTM.

We simulate N sample paths, following the construction in Section 3.3.2, and return the time elapsed for each path. To obtain the CDF, we define a set of points—which, when comparing to CDFs obtained via LST inversion is the same set of t values for which we invert the LSTs, referred to as `Tset` in Appendix B—and calculate the proportion of paths which terminate before t, for each value of t in our set.

We are also interested in the stability of the shifting procedure: time-shifting adds another layer of complexity to LST inversion, so where can things go wrong?

Throughout this section we use Algorithm B.3 to invert the LST of the TTM, and Algorithm B.4 for simulation, with 100,000 samples per CDF. We take the simulated CDF as the true CDF. More details about simulating the CDF are discussed in the above insert Simulated distributions.

We also assume from here on that the process starts in genus C at the upper boundary $b = 1$. This is the most sensible choice, since genus C tends to be the most favourable genus when the coral host is at full algal density.

4.4.1 Time scales

Figure 4.10 illustrates the CDF of the TTM at two time scales, obtained via LST inversion. We use the Euler and CME methods, with and without shifting, with $n = 11$ nodes. The simulated CDF is shown as a comparison.

There are several immediate observations to be made. From Figure 4.10a, we see that the CDF has a sudden jump at $t = 14.5$. This is a point mass, corresponding to the time taken for the level to drain from the upper boundary $b = 1$ to level $\zeta = 0.5$, plus the time taken for mortality to occur, τ. Recall that the maximum negative fluid rate is -1, and thus the former time is 0.5; adding $\tau = 14$ gives us our unstable point. From now on, we refer to this point as t_{crit}.

Another observation is that without shifting, the Euler and CME methods struggle to capture the timing of the jump at t_{crit}, whereas the shifted methods capture the timing of the jump exactly. This is also expected, since the LST of the

TTM has a shift of t_{crit} associated with it. Unfortunately, the value of the shifted CDFs after t_{crit} is slightly higher than the simulated value. This warrants further discussion, as the reason behind these higher values is not immediately clear.

Inaccurate shifting methods near t_{crit}

The inaccuracy of the shifting methods near t_{crit} is, unfortunately, unavoidable. Recall from Section 4.3.3 that the shifted methods produce larger values of the probability $\boldsymbol{e}_1^\intercal(\mathbf{1}-\boldsymbol{\Xi}^*\mathbf{1})$ than the simulated value, due to inaccuracies in computing the quantity $\boldsymbol{\Xi}^*$. The values of the CDF of the TTM immediately after t_{crit} are precisely $(e^{\zeta \mathbf{T}_{3,3}})\,\boldsymbol{e}_1^\intercal(\mathbf{1}-\boldsymbol{\Xi}^*\mathbf{1})$, where $\mathbf{T}_{3,3}$ is the diagonal entry of $\mathbf{T}$ corresponding to starting in phase C_-, since $e^{\zeta \mathbf{T}_{3,3}}$ is the probability of staying in phase C_- for the entire journey from level b to level ζ, and $\boldsymbol{e}_1^\intercal(\mathbf{1}-\boldsymbol{\Xi}^*\mathbf{1})$ is the probability of mortality before returning to ζ, starting in phase C_- and ending in any phase.

Thus, the quantity $\boldsymbol{\Xi}^*$ is the fundamental problem behind the inaccurate values of the CDF of the TTM near t_{crit}, when shifted methods are used.

4.4.2 Error analysis

So far, we have used the simulated CDF of the TTM to visually compare between the Euler and CME methods. A more robust method of comparison is to compare the CDFs with some error metric, which is what we consider now.

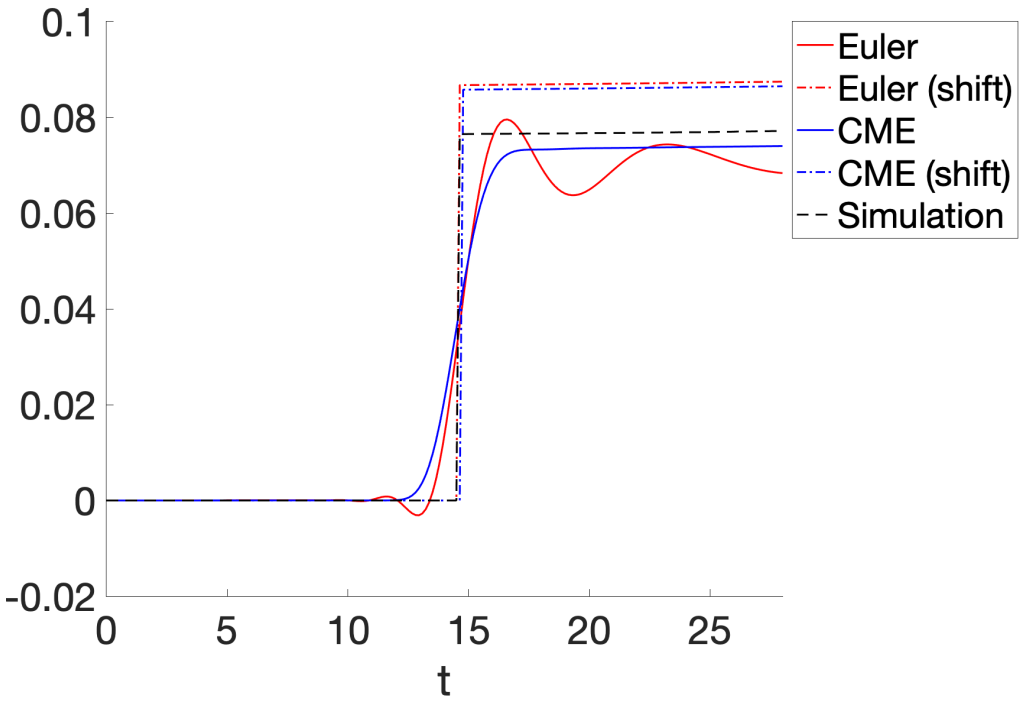
Recall that our simulated CDF is evaluated at the same points as the CDFs obtained via LST inversion. In addition, all of our CDFs are discretised—they are sequences of values, rather than actual functions. Thus, an appropriate error metric is the ℓ^2-norm, which is a distance measure for sequences.

For sequences $\boldsymbol{f} = \{f_k\}_{k=1}^K$ and $\boldsymbol{g} = \{g_k\}_{k=1}^K$, the ℓ^2-norm $\|\boldsymbol{f}-\boldsymbol{g}\|_2$ is

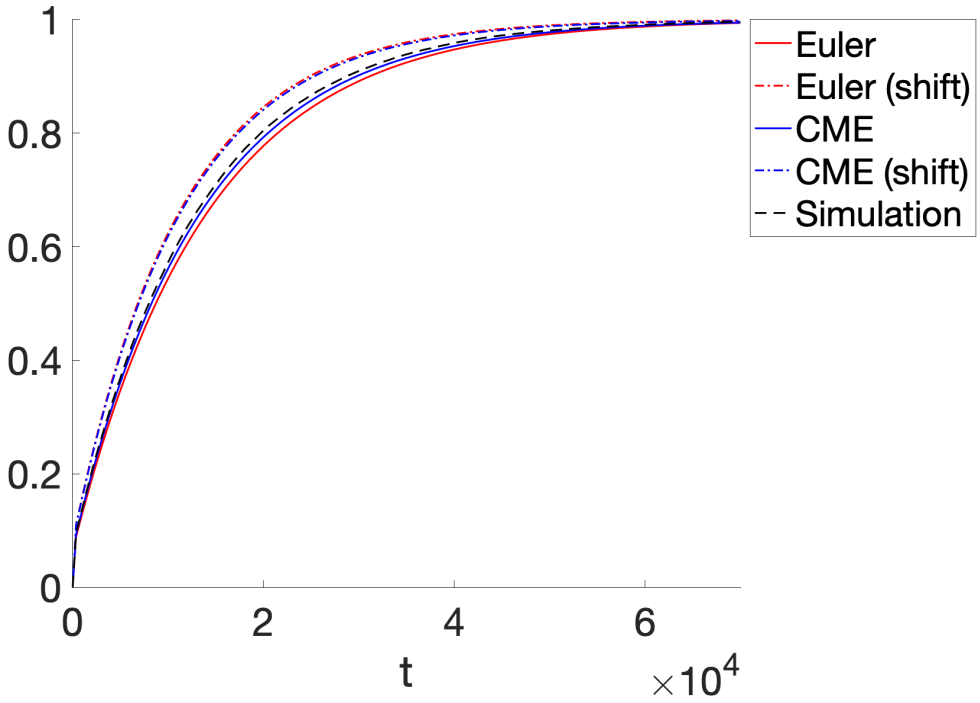
$$\|\boldsymbol{f}-\boldsymbol{g}\|_2 = \sqrt{\sum_{k=1}^{K} |f_k - g_k|^2}.$$

In our case, $\boldsymbol{f}$ is the sequence of CDF values under a specified method, and $\boldsymbol{g}$ is the sequence of simulated CDF values.

We would like to investigate how this error metric changes as we increase the number of nodes used in the LST inversion process. Intuitively, increasing the number of nodes should improve the approximation, since a larger number of nodes is equivalent to a higher order approximation to the Dirac delta functional in the Abate-Whitt framework.



(a) Short time scale. The shifted CDFs capture the jump at $t = 14.5$, but produce a higher probability of initial mortality.



(b) Long time scale. The shifted CDFs appear to be less accurate than their non-shifted counterparts.

Figure 4.10: CDF of the TTM, obtained via LST inversion, using the Euler and CME methods ($n = 11$ nodes) with and without shifting. Simulated CDF indicated by the dashed black line (using Algorithm B.4 with 100,000 samples).

Again, we investigate the behaviour of each method on two time scales, using $K = 200$ points to discretise the CDFs. The results are shown in Figure 4.11.

On both time scales the unshifted CME and Euler methods are more accurate according to the ℓ^2-norm, with the unshifted CME method almost always having the best performance. The unshifted Euler method appears to increase in error after $n = 31$ nodes, but this is likely due to the oscillations which typically get worse as n increases.

Interestingly, the shifted Euler and CME methods get worse as the number of nodes increases, although the shifted CME method appears to plateau after $n = 21$ nodes. This is likely to be due to the same phenomenon from Section 4.3.3, with the result that shifting cannot produce an accurate value very close to t_{crit}.

4.4.3 Shifting stability

We now pivot slightly, and investigate the stability of the inversion procedure; specifically, we are interested in the stability of atom-based shifting applied to the TTM. So far, the CME method without shifting has proven to be the most robust member of our suite of LST inversion methods, but this may change under different circumstances.

We focus on three ways for the inversion procedure to become unstable:

- Using a shortcut to compute $\int_0^\tau \boldsymbol{\xi}(u)\,du$.
- Using a large number of nodes.
- Using a set of parameters where the probability of mortality is small.

Numerical integration

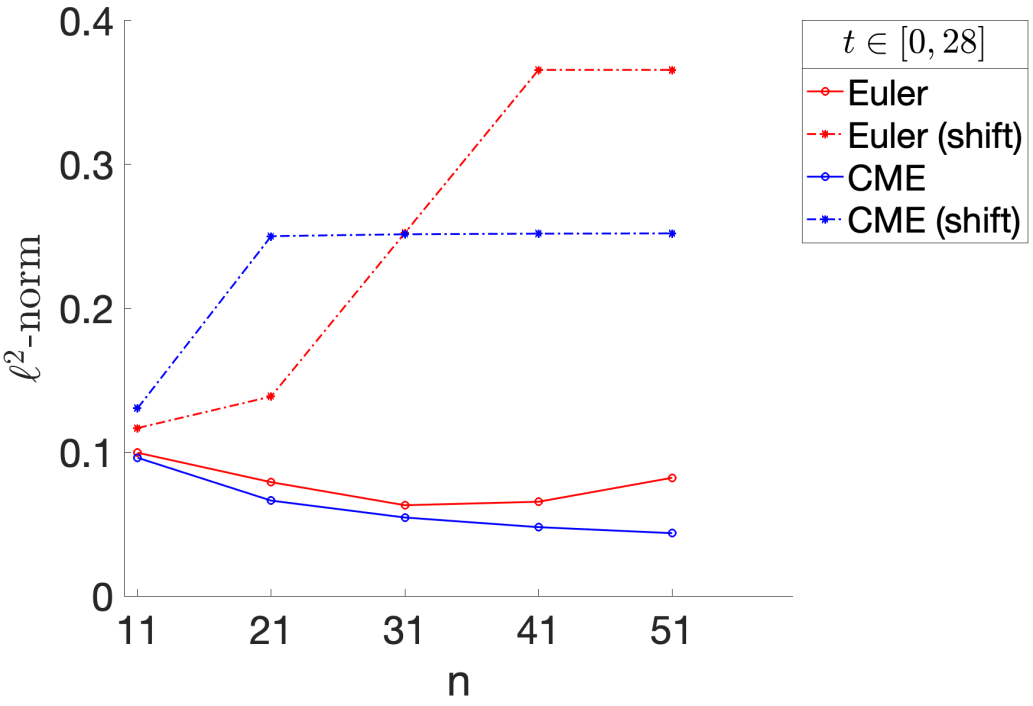
The second component of the TTM contains the integral expression

$$\int_0^\tau e^{-su}\boldsymbol{\xi}(u)\,du, \tag{4.5}$$

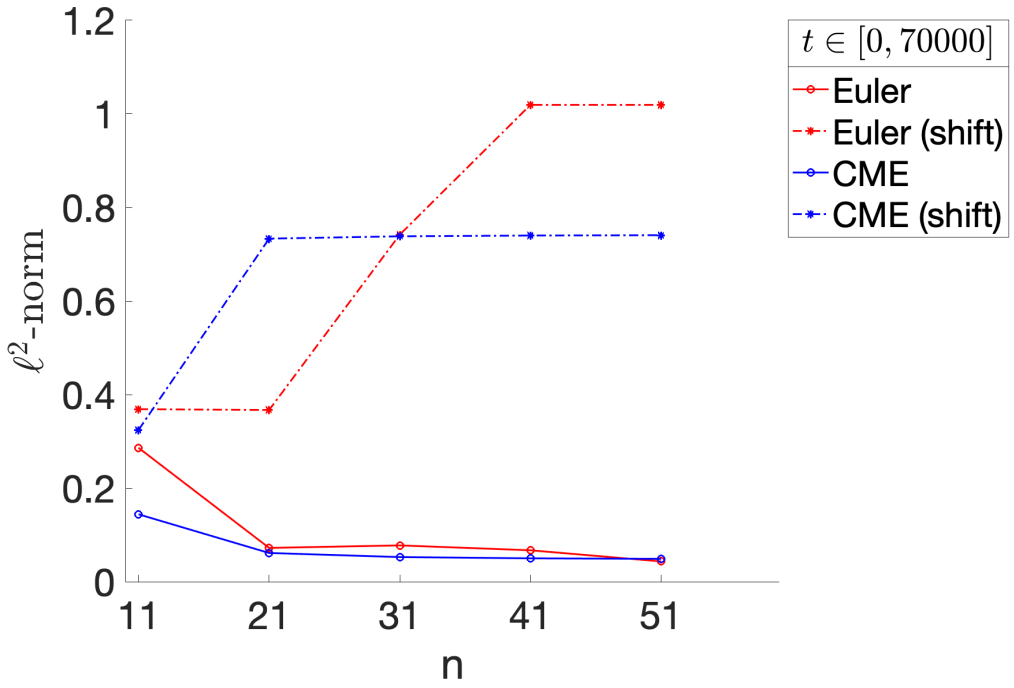
which represents the LST of the time spent below ζ before returning to ζ (at a time $t < \tau$). The third part of the TTM contains a similar integral:

$$e^{-s\tau}\left(\mathbf{1}_- - \boldsymbol{\Xi}^*\mathbf{1}_+\right) = e^{-s\tau}\left(\mathbf{1}_- - \int_0^\tau \boldsymbol{\xi}(u)\,du\,\mathbf{1}_+\right), \tag{4.6}$$

which represents the probability of mortality upon a journey below ζ, multiplied by the LST of the 'time cost' of this journey. This 'time cost' will always be τ, and thus the LST of this cost is the LST of $H_\tau(t)$, which is the LST $e^{-s\tau}$.



(a) ℓ^2-norm for short time scale: $t \in [0, 28]$.



(b) ℓ^2-norm for long time scale: $t \in [0, 70000]$.

Figure 4.11: ℓ^2-norm between the simulated CDF of the TTM (using Algorithm B.4 with 100,000 samples) and the CDF obtained via LST inversion, using the Euler and CME methods, with and without shifting. Number of nodes is on the horizontal axis, taking the values $n \in \{11, 21, 31, 41, 51\}$. Lines are shown to help illustrate trends. For each method, the ℓ^2-norm is calculated from 200 equally spaced points in the designated ranges.

In the evaluation of $F_{\text{TTM}}(t)$, both integrals must be computed numerically. A popular method—which we employ—is trapezoidal integration. This integration is taxing, since we are forced to obtain the density $\boldsymbol{\xi}(t)$ at a set of points (`Txi` in Algorithm B.3) to perform this trapezoidal integration.

It is unavoidable to evaluate Eq. (4.5) in this way, but it is possible to simplify the evaluation of Eq. (4.6). $\boldsymbol{\xi}(t)$ is the PDF of the time spent below ζ in the bounded model, but Eq. (4.6) is precisely the *CCDF* of this time, evaluated at $t = \tau$. Thus, instead of performing a second numerical integration, we can simply invert the LST of ${}_0^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)$ at $t = \tau$, making sure to convert to the CDF in the inversion (so we actually invert ${}_0^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)/s$), and then subtract the values from 1 to obtain the CCDF.

Great, we've evaded an *entire* integral!

The only problem is that this method does not hold up numerically, since it relies heavily on the accuracy of the CDF evaluated at τ. If this value is not correct, then this will affect the accuracy of the TTM, as illustrated in Figure 4.12a, where the CDFs exceed a value of 1 (or fall short of 1, in the case of the shifted CME method), which is very undesirable. Numerically evaluating the integral seems more preferable.

Large number of nodes

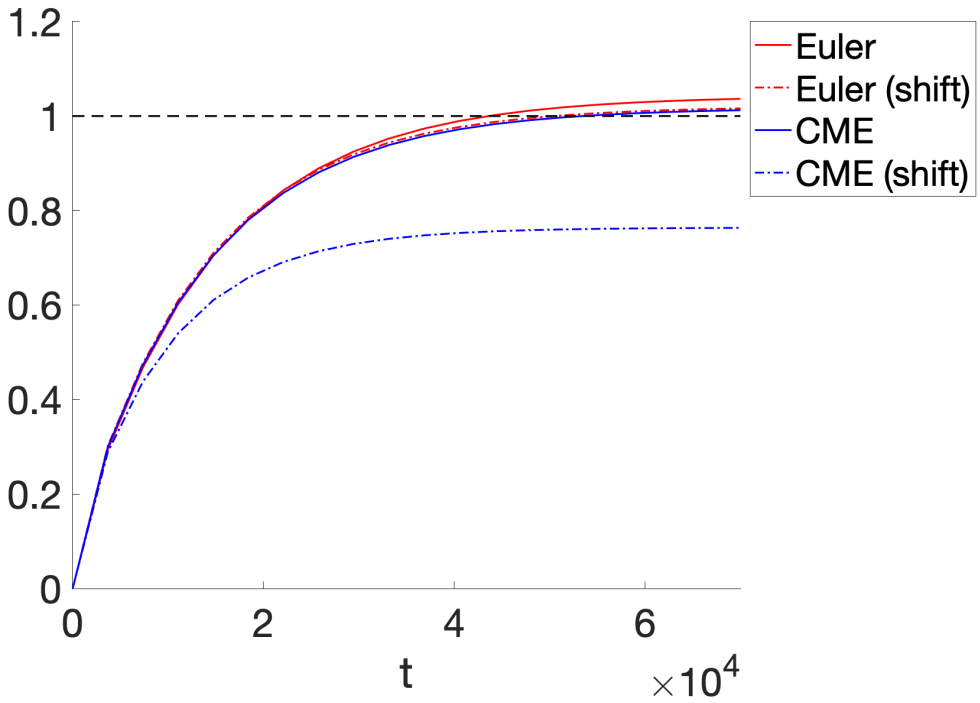
So far, we have only considered the inversion process for a relatively small number of nodes, up to $n = 51$. We might consider using a larger number of nodes to improve the accuracy of our CDFs; however, this can lead to numerical issues.

With $n = 201$ nodes, the Euler method (with and without shifting) completely breaks down. The inversion produces values as extreme as -9×10^{34}, and the inverse LST loses all resemblance to a CDF. The CME and shifted CME methods work perfectly fine with this many nodes.

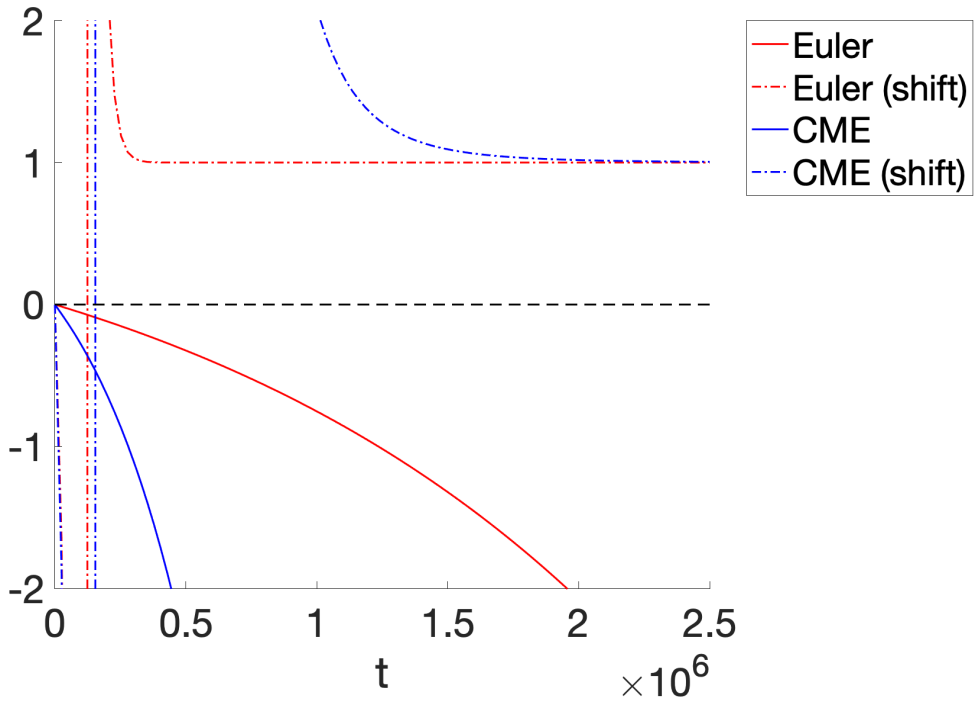
Small probability of mortality

Another source of numerical error occurs when the probability of mortality is very small in the coral fluid model. This requires a change in parameters, and thus no longer applies to the canonical model, but highlights another potential weakness of LST inversion.

Figure 4.12b illustrates this, created using a value of $\tau = 50$ (all other parameters are the canonical values), and we observe the chaos that plagues each of the CDFs. Strangely, the shifted CDFs seem to regain some stability after a certain amount of time, and plateau at 1—albeit, from above!



(a) CDF of the TTM, obtained via LST inversion with $n = 11$ nodes, where $\int_0^\tau \boldsymbol{\xi}(u)\,du$ is calculated by inverting ${}_0^{\zeta}\widetilde{\boldsymbol{\Xi}}(s)/s$ at $t = \tau$. Mathematically, this is correct, but numerically this causes issues.



(b) CDF of the TTM using a value of $\tau = 50$. CDFs obtained by LST inversion using $n = 11$ nodes. A value of $\tau = 50$ corresponds to a very low probability of mortality, since it takes a lot longer for the coral to die when the level hits ζ.

Figure 4.12: Issues that can arise when inverting the LST of the TTM.

This problem occurs due to the LST in Eq. (4.5). When τ is very large, or ζ is very small, the coral is much more likely to survive a journey below ζ. As a result, $\mathbf{I} - \int_0^\tau e^{-su}\boldsymbol{\xi}(u)\,du\,{}_\zeta^b\widetilde{\boldsymbol{\Psi}}(s)$ becomes a near-singular matrix, and thus, its inverse is the source of the problems observed in Figure 4.12b.

4.4.4 Interpreting the CDF

Consider a coral reef containing many coral colonies, each undergoing the same mortality process, under our fluid model. We are concerned with how many colonies are killed by a succession of bleaching events in a given time period.

Consider the problem on a short time scale. If t_{crit} is the first unstable point of $F_{\text{TTM}}(t)$, then $F_{\text{TTM}}(t_{\text{crit}})$ is the probability that a colony succumbs to mortality at the first possible time. On the reef scale, $F_{\text{TTM}}(t_{\text{crit}})$ is a proxy for the proportion of colonies killed almost immediately by the very first bleaching event.

Conversely, consider the problem on a longer time scale. Once again, we are presented with a CDF of the TTM; for some time $\bar{t}$, the value $F_{\text{TTM}}(\bar{t})$ is the probability that a colony has died before $\bar{t}$. On the reef scale, $F_{\text{TTM}}(\bar{t})$ is a proxy for the proportion of colonies which have died by $\bar{t}$, or alternatively, $1 - F_{\text{TTM}}(\bar{t})$ is a proxy for the proportion of colonies which are alive at $\bar{t}$. For example, if $\bar{t} = 5 \times 365$, then $1 - F_{\text{TTM}}(t)$ is the proportion of corals which survive 5 years of bleaching events. This is a *tail probability*.

To summarise:

- On a short time scale, the value of the CDF of the TTM at t_{crit} is a proxy for the proportion of coral colonies which immediately die from bleaching.
- On a long time scale, the *tail probabilities* are a proxy for the proportion of coral colonies which have survived up until a given time.

4.4.5 Method comparison

We are now in a position to compare the efficacy of the Euler and CME methods when applied to the TTM, and whether time-shifting is a valuable addition to both methods.

The Euler method (no shifting)

The biggest drawback of the unshifted Euler method is that it does not produce a proper CDF: it takes negative values immediately before t_{crit}, can exceed 1, and is often non-monotone. The inverse LST oscillates around t_{crit}, and these

oscillations worsen as the number of nodes, n, increases. This being said, even for a small number of nodes, it is very effective at capturing the long-term behaviour of the TTM, as evidenced in Figure 4.11b. Altogether, the method is able to produce a rough estimate of the distribution of the TTM, but is unsuitable for cases when a proper distribution is required, or when we are interested in the behaviour around t_{crit}.

The CME method (no shifting)

The unshifted CME method is arguably the most favourable out of the suite of methods we have considered. It has proven to be the most robust, and produce the most accurate estimate of the distribution of the TTM; by using a large number of nodes, we can produce a very accurate approximation.

The Euler method (shifting)

Although the shifted Euler method has proven to be the least accurate of the methods in most cases, it shows that time-shifting can fix some of the issues caused by the Euler method. The CDFs produced by the shifted Euler method do not have the adverse properties of the unshifted Euler method, and the method is able to capture the sharp jump at t_{crit} very nicely. Unfortunately, it underperforms due to the limitations of time-shifting.

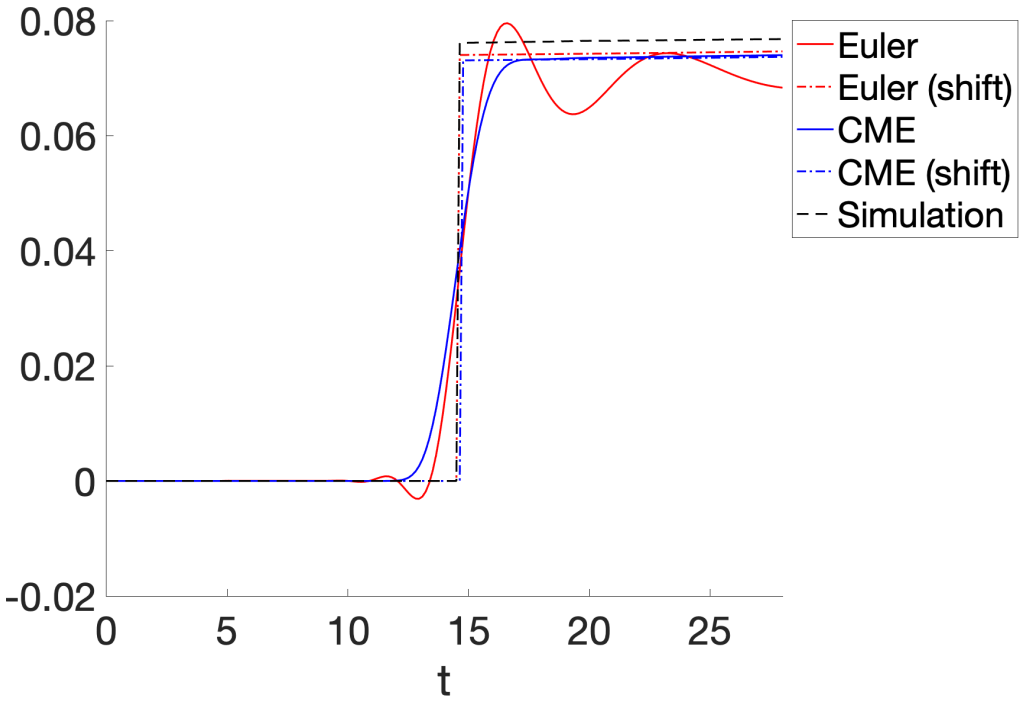
The CME method (shifting)

The shifted CME method also underperforms due to the limitations of time-shifting; however, it also illustrates that time-shifting helps capture sharp jumps for a low number of nodes. Using shifting with the CME method does not seem as valuable as with the Euler method since the unshifted CME method does very well on its own, but it may be useful for other applications.

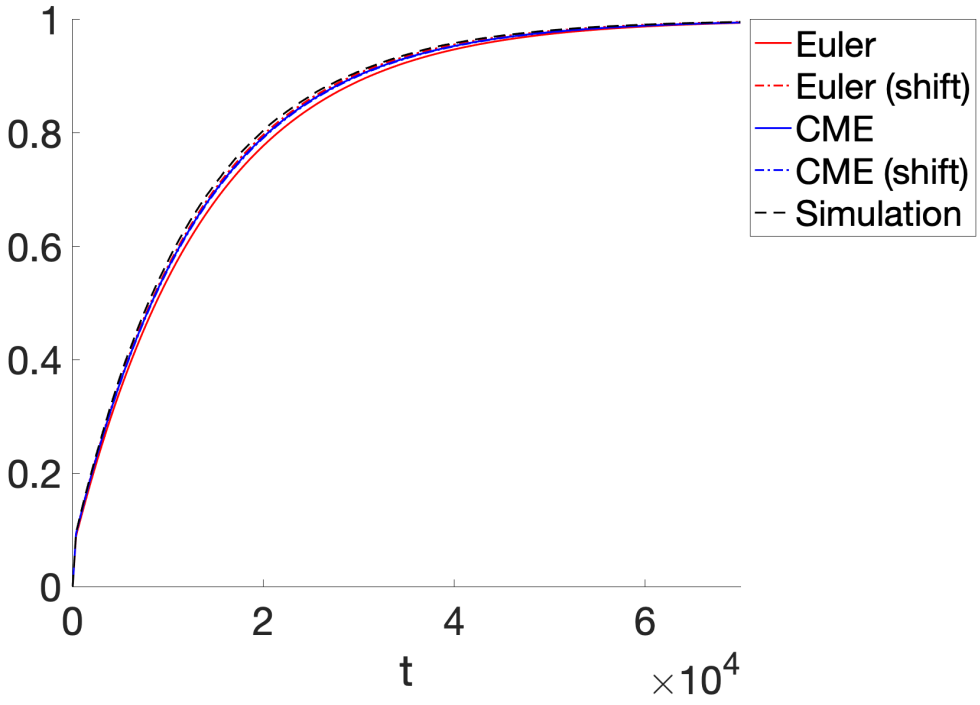
Improving upon shifting methods

Previously, we identified that shifting methods do not perform well very close to the point of shifting due to NaN values, which leads to an inaccurate value of $\boldsymbol{\Xi}^*$, and hence, an inaccurate value immediately after t_{crit}. To fix this, we could extrapolate $\boldsymbol{\xi}(t)$ where NaN values have been produced, in the hope of improving the approximation.

Figure 4.13 shows the CDF of the TTM where we extrapolate $\boldsymbol{\xi}(t)$ in a very simple manner: for all values $\boldsymbol{\xi}(t_p), \ldots, \boldsymbol{\xi}(t_q)$ that produce NaN values, we simply set $\boldsymbol{\xi}(t_i) = \boldsymbol{\xi}(t_{q+1})$, which is the first 'stable' value. The result is pleasing,



(a) Short time scale.



(b) Long time scale.

Figure 4.13: CDF of the TTM obtained via LST inversion with the Euler and CME methods, using $n = 11$ nodes. Shifted methods use a corrected version of $\boldsymbol{\xi}(t)$ which counteracts issues introduced by NaN values. Dashed black lines indicate the simulated CDF (using Algorithm B.4 with 100,000 samples). Compared to Figure 4.10, the shifted methods have significantly improved.

and shows that we can improve both shifting methods with such a correction. Moreover, we could use more sophisticated methods such as a linear or spline extrapolation to produce a more accurate estimate of $\boldsymbol{\xi}(t)$.

When applied to the TTM, this is very helpful, but in general, this approach should be applied on a case-by-case basis; whether correction is necessary, and how to implement it may not be clear. For an arbitrary LST, it may be unnecessary or detrimental to do this.

Back to coral

We may also think about each of our methods in terms of modelling coral-algal symbiosis. If we are interested in obtaining accurate tail probabilities, then the unshifted CME method seems to be the better choice, since it is the most accurate and robust in most cases.

If we are more interested in obtaining accurate estimates of the probability of immediate mortality, then perhaps the shifted Euler and shifted CME methods are more appropriate, since they are better at capturing the time of the sharp jump at t_{crit}. Moreover, with the aforementioned correction to $\boldsymbol{\xi}(t)$, these methods are able to capture the jump very accurately.

4.4.6 Summary

The results in this section do not shed all that much light on the processes of coral-algal symbiosis—and, as a result, may disappoint more biologically inclined readers—but do shed some light on the process of numerical LST inversion, particularly since the LST of the TTM is a relatively complex one.

Out of the Euler and CME methods, it seems clear that the CME method is, in general, superior in both accuracy and consistency when it comes to the TTM (although it is not entirely infallible, as shown in Figure 4.12b). As such, the CME method has proven itself as the favourable method for LST inversion.

Time-shifting has proven itself useful in some regard, but has several shortcomings, of which obtaining $\boldsymbol{\xi}(t)$ is the frontrunner. Although we can improve upon the shifting methods by correcting $\boldsymbol{\xi}(t)$ after performing this LST inversion, in a more general problem it may be difficult to detect if such an issue is present. It is also possible that the simplification discussed in Shifting disclaimer leads to some inaccuracies in time-shifting, which is an area for further investigation.

For simpler LSTs, time-shifting showed significant promise, and perhaps it is only due to the complexity of the TTM that it has been so troublesome. For simpler applications, it is a powerful technique that can greatly improve the Euler

method. The CME method is itself very powerful, but shifting can help it reach a higher level of accuracy without increasing the number of LST inversions.

The quantities $\boldsymbol{\xi}(t)$ and $\boldsymbol{\Xi}^*$ are highly significant in the inversion process. The embedding of another LST inversion appears to greatly complicate the overall process. In addition, there are several hidden parameters related to $\boldsymbol{\xi}(t)$ which affect the end result, such as the set of t values used to discretise $\boldsymbol{\xi}(t)$, the number of nodes used to obtain $\boldsymbol{\xi}(t)$, and even the inversion method used to obtain $\boldsymbol{\xi}(t)$, which could well be chosen to be different to the method used to obtain the distribution of the TTM.

The computational demand of the process should not be understated. Using atom-based shifting further complicates and adds additional computational complexity to this process. In addition, depending on the implementation, atom-based shifting can have severe performance issues, which was previously discussed in Atoms and molecules III.

Despite these issues, LST inversion is an efficient means to obtain interpretable time-dependent quantities in fluid models. Throughout this chapter, we used a simulated CDF for comparison, which, for 100,000 samples, can take over six times as long as inverting the LST at 200 points (for $n = 11$ nodes). Using shifting, simulation is approximately four times slower. Using less points or fewer nodes can increase the efficiency even further.

Towards a better model

At the end of Chapter 3, we drew two paths forward, one of which we have followed to conclusion. The second path goes in a new direction, and attempts to make our existing model more realistic biologically, and this is the path we follow next.

Chapter 5

Parisian Ruin and Erlangization

In this chapter, we explore an extension to the bounded fluid model considered in Chapter 3 and Chapter 4, with the aim of creating a model which is more biologically realistic. To do this, we explore the concept of Parisian ruin, borrowed from risk theory, which has a nice application in the coral-algal modelling framework. We create a new fluid model for coral-algal symbiosis with a Parisian-style definition of mortality, allowing for more versatility and biological realism, while remaining inside the existing fluid model framework. To achieve this, we explore a level-dependent model, and express the Parisian time to mortality (TTM) using familiar quantities from Chapter 3. These additional features make it difficult to obtain analytical results for the TTM; however, the model can still be simulated. To explore the performance of the model, we choose an Erlang horizon, and show that the distribution of the TTM in the 'Erlangized' model converges to the distribution of the TTM under the basic model from Chapter 3 and Chapter 4. This agrees with an existing result which shows that the probability of mortality in the Erlangized model converges to the probability in the basic model, although the convergence in distribution is not monotonic, as might be expected. The Parisian model is particularly important due to its relation to the model presented in Chapter 6, which further extends the concepts explored in this chapter.

5.1 Parisian ruin

In addition to sounding groovy, the phase process of a Markov-modulated fluid model allows for the incorporation of incredibly complex behaviour: one might even say it's the powerhouse of the model. In this section, we introduce the notion of a *horizon*: a random time at which our fluid process is terminated. In our model for coral-algal symbiosis, mortality occurs when the horizon is reached.

The idea is to incorporate this horizon into the phase process itself, allowing for a fluid model where coral mortality occurs at a *random* time, rather than a fixed time τ, whilst remaining in the existing fluid model framework. Therefore, our model attains a higher level of functionality, but we can still use many of the analytical techniques already seen in this thesis.

5.1.1 Fluid models with a horizon

The notion of adding a phase-type (PH) distributed horizon to a fluid model was first explored in the context of risk processes in Asmussen, Avram and Usabel [4].

Peralta [66] extends this notion to obtain the probability of Parisian and cumulative Parisian ruin in a standard risk model.

In risk analysis, ruin is typically defined as the level of a process (e.g., the amount of money in a bank account) falling below 0. Parisian ruin occurs when the process does not recover before the expiry of some randomly-distributed clock after falling below level 0, where the clock is reset upon each return to level 0. Cumulative Parisian ruin is identical, but the random clock is never reset.

There is an obvious parallel here to coral mortality, which we have previously defined as the fluid process spending τ time below a critical threshold ζ, and thus a sensible extension to our model is to consider a Parisian-style definition of mortality, where the time τ, now *random*, has a PH distribution.

Our approach will be different to that of the above references in several ways. We are primarily interested in time-dependent quantities, whereas previous approaches have only considered probabilities. Additionally, we consider a bounded model, whereas previous approaches have only considered unbounded fluid models.

5.1.2 The augmented model

Here we define a fluid process augmented with a continuous PH distributed horizon, or clock. In this model, which we refer to as the *augmented model*, we define mortality to be the time at which the PH clock expires. The aim is to once again obtain the distribution of the time to mortality (TTM) under this model by use of Laplace-Stieltjes transforms (LSTs).

The new model will be constructed from the existing unbounded fluid process $\{\mathcal{F}(t)\}_{t\geq 0}$ from Section 2.4, defined by the matrices $\mathbf{T}$ and $\mathbf{C}$.

Define the horizon H as a phase-type random variable on the state space $\mathcal{S}_\ell = \{1, 2, \ldots, \ell\}$ with a *cemetery* (absorbing) state 0. We say $H \sim \mathrm{PH}_\ell(\boldsymbol{\kappa}^\intercal, \mathbf{K})$, where $\boldsymbol{\kappa}^\intercal$ is the initial state distribution and $\mathbf{K}$ is the $\ell \times \ell$ transient generator of the CTMC representing the PH distribution. Let us denote by $\boldsymbol{k} = -\mathbf{K}\mathbf{1}_\ell$ the

vector of rates from $\mathcal{S}_\ell$ into the cemetery state 0. We thus define the augmented (unbounded) fluid model $\{\mathcal{F}^\#(t)\}_{t\geq 0} = \{\big(M^\#(t), \varphi^\#(t)\big) : t \geq 0\}$ as follows:

- The phase process $\varphi^\#(t)$ is a CTMC on the state space $\mathcal{S}^\# \cup \{0\}$, where $\mathcal{S}^\# = \mathcal{S} \times \mathcal{S}_\ell$. We order $\mathcal{S}^\#$ lexicographically: $(i_1, i_2) \leq (j_1, j_2) \iff i_1 < j_1$ or $i_1 = j_1, i_2 \leq j_2$. For convenience, let $s^\# = |\mathcal{S}^\#| = |\mathcal{S}| \times \ell$.

- When $\varphi^\#(t) \in \mathcal{S}^\#$, the process evolves according to the infinitesimal generator $\mathbf{T}^\#$. When $\varphi^\#(t) = 0$ (the cemetery state), the horizon is reached and the process terminates.

- When $\varphi^\#(t) = (i_1, i_2) = \boldsymbol{i} \in \mathcal{S}^\#$, the level process $M^\#(t)$ changes at rate c_{i_1}. These rates are stored in the matrix $\mathbf{C}^\# = \mathbf{C} \otimes \mathbf{I}_\ell$, where $[\mathbf{C}^\#]_{\boldsymbol{i},\boldsymbol{i}} = |c_{i_1}|$.

- We partition $\mathcal{S}$ in the usual way: $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_- \cup \mathcal{S}_0$. From here on, we assume that $\mathcal{S}_0 = \emptyset = \mathcal{S}_0^\#$, and subsequently $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_-$. We partition $\mathcal{S}^\#$ in the exact same way: $\mathcal{S}^\# = \mathcal{S}_+^\# \cup \mathcal{S}_-^\#$, where $\mathcal{S}_+^\# = \mathcal{S}_+ \times \mathcal{S}_\ell$ and $\mathcal{S}_-^\# = \mathcal{S}_- \times \mathcal{S}_\ell$. In addition, let $s_+^\# = |\mathcal{S}_+^\#|$ and $s_-^\# = |\mathcal{S}_-^\#|$.

- The matrix $\mathbf{T}^\#$ can thus be written as

$$\mathbf{T}^\# = \begin{bmatrix} \mathbf{T}_{++}^\# & \mathbf{T}_{+-}^\# \\ \mathbf{T}_{-+}^\# & \mathbf{T}_{--}^\# \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{++} \oplus \mathbf{K} & \mathbf{T}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}_{--} \oplus \mathbf{K} \end{bmatrix} = \mathbf{T} \oplus \mathbf{K}.$$

  Note that $\mathbf{T}^\#$ is not conservative: in some states, there is a positive rate of exiting $\mathcal{S}^\#$ (these rates are in the vector $-\mathbf{T}^\#\mathbf{1}$), since H is finite with positive probability. If such a transition occurs, the phase process transitions into the cemetery state 0. Subsequently, $\mathbf{K}$ is not conservative either.

  The structure of $\mathbf{T}^\#$ has a useful interpretation:

  - First, observe that $\mathbf{T}^\# = \mathbf{T} \oplus \mathbf{K}$ can be written as

$$\begin{bmatrix} \mathbf{T}_{++} \oplus \mathbf{K} & \mathbf{T}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}_{--} \oplus \mathbf{K} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}_{--} \otimes \mathbf{I}_\ell \end{bmatrix} + \begin{bmatrix} \mathbf{I}_+ \otimes \mathbf{K} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_- \otimes \mathbf{K} \end{bmatrix}.$$

  - The matrices $\mathbf{T}_{++}$, $\mathbf{T}_{+-}$, and so on operate on the state space $\mathcal{S}$; taking the Kronecker product with $\mathbf{I}_\ell$ allows the new matrices $\mathbf{T}_{++}^\#$, $\mathbf{T}_{+-}^\#$, and so on to operate on the state space $\mathcal{S}^\#$.

  - The matrices $\mathbf{T}_{++}^\#$ and $\mathbf{T}_{--}^\#$ represent transitions in either the phase process *or* the horizon process, but not both. For example, consider the submatrix $\mathbf{T}_{++}^\# = \mathbf{T}_{++} \oplus \mathbf{K} = \mathbf{T}_{++} \otimes \mathbf{I}_\ell + \mathbf{I}_+ \otimes \mathbf{K}$: the matrix

$\mathbf{T}_{++} \otimes \mathbf{I}_\ell$ represents transitions in the phase process (within $\mathcal{S}_+$) where nothing occurs in the horizon process; the matrix $\mathbf{I}_+ \otimes \mathbf{K}$ represents transitions in the horizon process where nothing happens in the phase process. The submatrix $\mathbf{T}^{\#}_{--}$ can be interpreted similarly.

– The matrices $\mathbf{T}^{\#}_{+-}$ and $\mathbf{T}^{\#}_{-+}$ represent transitions in the phase process only (from $\mathcal{S}_+$ to $\mathcal{S}_-$ or vice-versa), since it is impossible for the horizon to change at the same time as the phase.

In essence, the augmented process $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$ is almost identical to the original process $\{\mathcal{F}(t)\}_{t\geq 0}$, but with an independent horizon process attached to it.

$\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$ is analogous to the risk process $\{R^*_t\}_{t\geq 0}$ defined in Peralta [66], which was argued to have the same first passage probabilities as the simpler process $\{R_t\}_{t\geq 0}$ (analogous to $\{\mathcal{F}(t)\}_{t\geq 0}$) terminated when the horizon is reached. We can similarly conclude that the process $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$ has the same first passage probabilities as the process $\{\mathcal{F}(t)\}_{t\geq 0}$ terminated at time H. Furthermore, if $\{\mathcal{F}(t)\}_{t\geq 0}$ is terminated at H, then this process taboos any paths which take a time exceeding H, which is precisely the process $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$: thus, $\{\mathcal{F}(t)\}_{t\geq 0}$ and $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$ have the same first passage *times*, in addition to probabilities.

5.1.3 Quantities in the augmented model

Probabilities

Using the same techniques as in Section 2.4, we are able to derive important probabilities such as the matrices $\mathbf{\Psi}^{\#}$ and $\mathbf{\Xi}^{\#}$, which give the probability of return to initial level (before the horizon is reached), starting from an upwards or downwards phase, respectively. For $\boldsymbol{i} = (i_1, i_2) \in \mathcal{S}^{\#}_+$ and $\boldsymbol{j} = (j_1, j_2) \in \mathcal{S}^{\#}_-$,

$$\left[\mathbf{\Psi}^{\#}\right]_{\boldsymbol{i},\boldsymbol{j}} = \mathbb{P}\left[\theta^{\#}(z) < H, \varphi^{\#}(\theta^{\#}(z)) = \boldsymbol{j} \;\middle|\; M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{i}\right],$$
$$\left[\mathbf{\Xi}^{\#}\right]_{\boldsymbol{j},\boldsymbol{i}} = \mathbb{P}\left[\theta^{\#}(z) < H, \varphi^{\#}(\theta^{\#}(z)) = \boldsymbol{i} \;\middle|\; M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{j}\right],$$

where $\theta^{\#}(x) = \inf\{t > 0 : M^{\#}(t) = x\}$ is the first passage time to level x. As with $\mathbf{\Psi}$ and $\mathbf{\Xi}$, these can be obtained by solving algebraic Riccati equations similar to Eq. (2.11) and Eq. (2.13), respectively.

Peralta [66] obtains the probability of Parisian ruin, $\psi^F_i(u)$, in Section 5.2, by using the matrix $\boldsymbol{\alpha}^{*c}_{-+}$, equivalent to the $s^{\#}_- \times s^{\#}_+$ matrix $\mathbf{\Xi}^{\#}$. $\psi^F_i(u)$ is the $s_+ \times 1$ probability that the horizon is reached, given that the process starts at level $u > 0$ in phase $i \in \mathcal{S}_+$, and that the horizon only progresses when the level is below 0.

Time-dependent quantities

We would like to extend the expression for the probability of Parisian ruin to obtain the distribution of time until Parisian ruin occurs. Following the structure of Section 2.4, we construct the building blocks for obtaining this distribution, beginning with the important quantities $\widehat{\boldsymbol{\Psi}}^{\#}(s)$ and $\widehat{\boldsymbol{\Xi}}^{\#}(s)$.

For $\boldsymbol{i} = (i_1, i_2) \in \mathcal{S}_+^{\#}$ and $\boldsymbol{j} = (j_1, j_2) \in \mathcal{S}_-^{\#}$, define the LSTs

$$\begin{aligned}\left[\widehat{\boldsymbol{\Psi}}^{\#}(s)\right]_{\boldsymbol{i},\boldsymbol{j}} &= \mathbb{E}\left[e^{-s\theta^{\#}(z)}; \theta^{\#}(z) < H, \varphi^{\#}(\theta^{\#}(z)) = \boldsymbol{j} \;\middle|\; M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{i}\right], \\ \left[\widehat{\boldsymbol{\Xi}}^{\#}(s)\right]_{\boldsymbol{j},\boldsymbol{i}} &= \mathbb{E}\left[e^{-s\theta^{\#}(z)}; \theta^{\#}(z) < H, \varphi^{\#}(\theta^{\#}(z)) = \boldsymbol{i} \;\middle|\; M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{j}\right],\end{aligned}$$

which are the LSTs of the time taken to return to the initial level z, and to do so in taboo of reaching the horizon H, starting from an upwards ($\widehat{\boldsymbol{\Psi}}^{\#}(s)$) or downwards ($\widehat{\boldsymbol{\Xi}}^{\#}(s)$) phase in the augmented model. We can obtain these LSTs via the following theorem.

Theorem 5.1. *$\widehat{\boldsymbol{\Psi}}^{\#}(s)$ solves the algebraic Riccati equation*

$$\mathbf{Q}_{+-}^{\#}(s) + \widehat{\boldsymbol{\Psi}}^{\#}(s)\mathbf{Q}_{-+}^{\#}(s)\widehat{\boldsymbol{\Psi}}^{\#}(s) + \mathbf{Q}_{++}^{\#}(s)\widehat{\boldsymbol{\Psi}}^{\#}(s) + \widehat{\boldsymbol{\Psi}}^{\#}(s)\mathbf{Q}_{--}^{\#}(s) = \mathbf{0}, \tag{5.1}$$

where

$$\begin{aligned}\mathbf{Q}_{++}^{\#}(s) &= \left(\mathbf{C}_+^{\#}\right)^{-1}\left[\mathbf{T}_{++}^{\#} - s\mathbf{I}\right], \\ \mathbf{Q}_{--}^{\#}(s) &= \left(\mathbf{C}_-^{\#}\right)^{-1}\left[\mathbf{T}_{--}^{\#} - s\mathbf{I}\right], \\ \mathbf{Q}_{+-}^{\#}(s) &= \left(\mathbf{C}_+^{\#}\right)^{-1}\mathbf{T}_{+-}^{\#}, \\ \mathbf{Q}_{-+}^{\#}(s) &= \left(\mathbf{C}_-^{\#}\right)^{-1}\mathbf{T}_{-+}^{\#}.\end{aligned} \tag{5.2}$$

Similarly, $\widehat{\boldsymbol{\Xi}}^{\#}(s)$ solves the symmetric Riccati equation with $+$ and $-$ switched.

If s is real, then each quantity is the minimal non-negative solution to its respective equation.

Proof. The result follows directly from Theorem 1 of Bean *et al.* [11]. The only departure—apart from the omission of $\mathcal{S}_0$—from this theorem is that $\mathbf{T}^{\#}$ is not conservative; this means we must only consider paths which return to the initial level in taboo of reaching the horizon H (this is explained in more detail in the insert Killed hitting times over the page). The remainder of the proof is the same. ■

Killed hitting times

The definitions of $\widehat{\boldsymbol{\Psi}}^{\#}(s)$ and $\widehat{\boldsymbol{\Xi}}^{\#}(s)$ make it is easier to see why the augmented process must have the same hitting times as the simple process killed at time H. In the original model, if the first passage time $\theta(z)$ occurs before the horizon is reached, then a path that returns to z takes the same amount of time as a path in the augmented model (which reaches z at $\theta^{\#}(z) = \theta(z)$).

If the first passage time $\theta(z)$ occurs after the horizon is reached, then that path can never return to z, since such a path is under taboo. This is also what happens in the augmented model, in which the taboo is manifested as the cemetery state 0.

With $\widehat{\boldsymbol{\Psi}}^{\#}(s)$ and $\widehat{\boldsymbol{\Xi}}^{\#}(s)$ we can derive the LSTs of filling and draining times, $\widehat{\mathbf{G}}^{\#}(s;z)$ and $\widehat{\mathbf{H}}^{\#}(s;z)$, which, starting in phase $\boldsymbol{i} \in \mathcal{S}^{\#}$ and ending in phase $\boldsymbol{j} \in \mathcal{S}^{\#}$, are defined as

$$\left[\widehat{\mathbf{G}}^{\#}(s;z)\right]_{\boldsymbol{i},\boldsymbol{j}} = \mathbb{E}\left[e^{-s\theta^{\#}(0)}\mathbb{I}\big(\theta^{\#}(0) < H, \varphi^{\#}(\theta^{\#}(0)) = \boldsymbol{j}\big) \;\middle|\; M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{i}\right],$$
$$\left[\widehat{\mathbf{H}}^{\#}(s;z)\right]_{\boldsymbol{i},\boldsymbol{j}} = \mathbb{E}\left[e^{-s\theta^{\#}(z)}\mathbb{I}\big(\theta^{\#}(z) < H, \varphi^{\#}(\theta^{\#}(z)) = \boldsymbol{j}\big) \;\middle|\; M^{\#}(0) = 0, \varphi^{\#}(0) = \boldsymbol{i}\right].$$

As usual, we can partition these quantities according to $\mathcal{S}_+^{\#}$ and $\mathcal{S}_-^{\#}$. We obtain $\widehat{\mathbf{G}}^{\#}(s;z)$ and $\widehat{\mathbf{H}}^{\#}(s;z)$ with the following theorem.

Theorem 5.2. *In the process $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$, for any $z > 0$, the respective LSTs of the time to drain from level z to 0 ($\widehat{\mathbf{G}}^{\#}(s;z)$) or fill the buffer from level 0 to z ($\widehat{\mathbf{H}}^{\#}(s;z)$), in taboo of paths in which the horizon H is reached before the destination is reached, are given by*

$$\begin{aligned}
\widehat{\mathbf{G}}^{\#}_{+-}(s;z) &= \widehat{\boldsymbol{\Psi}}^{\#}(s)\exp(\mathbf{B}^{\#}(s)z),\\
\widehat{\mathbf{G}}^{\#}_{--}(s;z) &= \exp(\mathbf{B}^{\#}(s)z),\\
\widehat{\mathbf{H}}^{\#}_{++}(s;z) &= \exp(\mathbf{A}^{\#}(s)z),\\
\widehat{\mathbf{H}}^{\#}_{-+}(s;z) &= \widehat{\boldsymbol{\Xi}}^{\#}(s)\exp(\mathbf{A}^{\#}(s)z),
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A}^{\#}(s) &= \left(\mathbf{C}^{\#}_{+}\right)^{-1}\left[\left(\mathbf{T}^{\#}_{++} - s\mathbf{I}\right) + \mathbf{T}^{\#}_{+-}\,\widehat{\boldsymbol{\Xi}}^{\#}(s)\right],\\
\mathbf{B}^{\#}(s) &= \left(\mathbf{C}^{\#}_{-}\right)^{-1}\left[\left(\mathbf{T}^{\#}_{--} - s\mathbf{I}\right) + \mathbf{T}^{\#}_{-+}\,\widehat{\boldsymbol{\Psi}}^{\#}(s)\right].
\end{aligned}$$

Proof. The result follows directly from Theorem 1 of Bean *et al.* [13]. Again, the only difference is that $\mathbf{T}^{\#}$ is not conservative, and thus we must only consider paths which reach level z in taboo of reaching the horizon H. The remainder of the proof is the same.

■

To complete our suite of unbounded quantities, we define the LSTs ${}_{x}^{y}\widehat{\mathbf{G}}^{\#}(s;z)$ and ${}_{x}^{y}\widehat{\mathbf{H}}^{\#}(s;z)$ as follows. Starting in $\boldsymbol{i} \in \mathcal{S}^{\#}$ and ending in phase $\boldsymbol{j} \in \mathcal{S}^{\#}$, these quantities are defined as

$$\left[{}_{x}^{y}\widehat{\mathbf{G}}^{\#}(s;z)\right]_{\boldsymbol{i},\boldsymbol{j}} = \mathbb{E}\left[e^{-s\theta^{\#}(x)}\mathbb{I}\big(\theta^{\#}(x) < \theta^{\#}(y), \theta^{\#}(x) < H, \varphi^{\#}(\theta(x)) = \boldsymbol{j}\big) \,\middle|\, \ldots \right.$$
$$\left. M^{\#}(0) = z, \varphi(0) = \boldsymbol{i}\right],$$
$$\left[{}_{x}^{y}\widehat{\mathbf{H}}^{\#}(s;z)\right]_{\boldsymbol{i},\boldsymbol{j}} = \mathbb{E}\left[e^{-s\theta^{\#}(y)}\mathbb{I}\big(\theta^{\#}(y) < \theta^{\#}(x), \theta^{\#}(y) < H, \varphi^{\#}(\theta^{\#}(y)) = \boldsymbol{j}\big) \,\middle|\, \ldots \right.$$
$$\left. M^{\#}(0) = z, \varphi^{\#}(0) = \boldsymbol{i}\right].$$

At this point, it should be clear that the same argument from Theorem 5.1 and Theorem 5.2 will apply to ${}_{x}^{y}\widehat{\mathbf{G}}^{\#}(s;z)$, ${}_{x}^{y}\widehat{\mathbf{H}}^{\#}(s;z)$ and any other fluid quantities we define in the augmented model. In addition, the same reasoning holds for a *bounded* augmented fluid model, which we define in the next section.

5.2 Bounded, level-dependent Parisian ruin

Here, we define the *Parisian model*: a bounded, level dependent version of the augmented model. The idea is to create two models: one for when the level is above ζ, and the horizon is not present; one for when the level is below ζ, and the horizon can progress. By 'stitching' the models together, we obtain a process with level-dependent behaviour. Moreover, this model has a very similar definition of mortality to the model in Chapter 3, where mortality is defined as spending too long below the threshold ζ; however, now the maximum time spent below ζ before mortality, τ, has a PH distribution, rather than being a constant.

Define the Parisian model $\{\widetilde{\mathcal{F}}^{\#}(t)\}_{t\geq 0} = \{(\widetilde{M}^{\#}(t), \widetilde{\varphi}^{\#}(t)) : t \geq 0\}$ as follows:

- When the level $\widetilde{M}^{\#}(t)$ is in the interval (ζ, b), the phase process $\widetilde{\varphi}^{\#}(t)$ evolves exactly like $\widetilde{\varphi}(t)$ with infinitesimal generator $\mathbf{T}$ on the state space $\mathcal{S}$. The level process $\widetilde{M}^{\#}(t)$ evolves exactly like $\widetilde{M}(t)$, with fluid input rates in $\mathbf{C}$.

- When the level $\widetilde{M}^{\#}(t)$ hits the upper boundary b, the phase process immediately enters $\widehat{\mathcal{S}}_0 \cup \mathcal{S}_-$ according to $\widehat{\mathbf{P}}$. While the phase is in $\widehat{\mathcal{S}}_0$, the fluid

input rate is 0 and the phase evolves according to $\widehat{\mathbf{T}}$ until it enters a phase in $\mathcal{S}_-$ (exactly as in the regular bounded process).

- When the level $\widetilde{M}^{\#}(t)$ is in the interval $(0, \zeta)$, the phase process $\widetilde{\varphi}^{\#}(t)$ evolves according to the infinitesimal generator $\mathbf{T}^{\#}$ on the state space $\mathcal{S}^{\#}$, exactly like $\varphi^{\#}(t)$ from $\{\mathcal{F}^{\#}(t)\}_{t\geq 0}$.

  The level process $\widetilde{M}^{\#}(t)$ evolves with fluid input rates in $\mathbf{C}^{\#} = \mathbf{C} \otimes \mathbf{I}_\ell$. Both $\mathbf{T}^{\#}$ and $\mathbf{C}^{\#}$ have the same definitions as in Section 5.1.2. If the PH clock expires (the cemetery state is reached), the process terminates.

- When the level $\widetilde{M}^{\#}(t)$ hits the lower boundary 0, the phase process immediately enters $(\check{\mathcal{S}}_0 \times \mathcal{S}_\ell) \cup (\mathcal{S}_+ \times \mathcal{S}_\ell)$ according to $\check{\mathbf{P}}^{\#}$. While the phase is in $\widehat{\mathcal{S}}_0 \times \mathcal{S}_\ell$, the fluid input rate is 0 and the phase evolves according to the infinitesimal generator $\check{\mathbf{T}}^{\#}$ until it enters a phase in $\mathcal{S}_+$. Here,

$$\check{\mathbf{P}}^{\#} = \begin{bmatrix} \check{\mathbf{P}}^{\#}_{-0} & \vdots & \check{\mathbf{P}}^{\#}_{-+} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{P}}_{-0} \otimes \mathbf{I}_\ell & \vdots & \check{\mathbf{P}}_{-+} \otimes \mathbf{I}_\ell \end{bmatrix},$$
$$\check{\mathbf{T}}^{\#} = \begin{bmatrix} \check{\mathbf{T}}^{\#}_{00} & \vdots & \check{\mathbf{T}}^{\#}_{0+} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{T}}_{00} \oplus \mathbf{K} & \vdots & \check{\mathbf{T}}_{0+} \otimes \mathbf{I}_\ell \end{bmatrix}.$$

- When the level $\widetilde{M}^{\#}(t)$ hits ζ from above in phase $i \in \mathcal{S}_-$, the phase process enters $(j_1, j_2) \in \mathcal{S}_- \times \mathcal{S}_\ell$ according to the $(i, (j_1, j_2))$ entry of $\check{\mathbf{P}}_\zeta = \mathbf{I}_- \otimes \boldsymbol{\kappa}^\intercal$.
- When the level $\widetilde{M}^{\#}(t)$ hits ζ from below in phase $(i_1, i_2) \in \mathcal{S}_+ \times \mathcal{S}_\ell$, the phase process enters $j \in \mathcal{S}_+$ according to the $((i_1, i_2), j)$ entry of $\widehat{\mathbf{P}}_\zeta = \mathbf{I}_+ \otimes \mathbf{1}_\ell$.

The behaviour of the Parisian model at level ζ allows for the 'stitching' of the upper and lower models together. When the level hits ζ from above, the phase process must transition from $\mathcal{S}_-$ into $\mathcal{S}_-^{\#}$. To do this, we multiply by $\check{\mathbf{P}}_\zeta = \mathbf{I}_- \otimes \boldsymbol{\kappa}^\intercal$. The multiplication by $\mathbf{I}_-$ keeps the information of the current phase: if the process is in $i \in \mathcal{S}_-$ above the threshold, it will be in $(i_1, i_2) \in \mathcal{S}_-^{\#}$ once the level crosses the threshold, which is the corresponding phase in the larger phase space. The index i_2 corresponds to the initial horizon stage, which is determined by $\boldsymbol{\kappa}^\intercal$.

Similarly, when the level hits ζ from below, the phase process must transition from $\mathcal{S}_+^{\#}$ into $\mathcal{S}_+$. Using the definition of mortality from Chapter 3, we would like the mortality process to reset once the level returns to the critical threshold. To achieve this, we can multiply by $\widehat{\mathbf{P}}_\zeta = \mathbf{I}_+ \otimes \mathbf{1}_\ell$. Once again, the multiplication by $\mathbf{I}_+$ keeps the information of the current phase: if the process is in $(i, j) \in \mathcal{S}_+^{\#}$ below the threshold, it will be in $i \in \mathcal{S}_+$ once the level crosses the threshold. The horizon stage j is subsequently lost, which captures the idea that mortality 'resets' once the process returns to ζ. This is captured by the $\mathbf{1}_\ell$ term in $\widehat{\mathbf{P}}_\zeta$.

The behaviour of the phase process at level ζ is further illustrated in Figure 5.1.

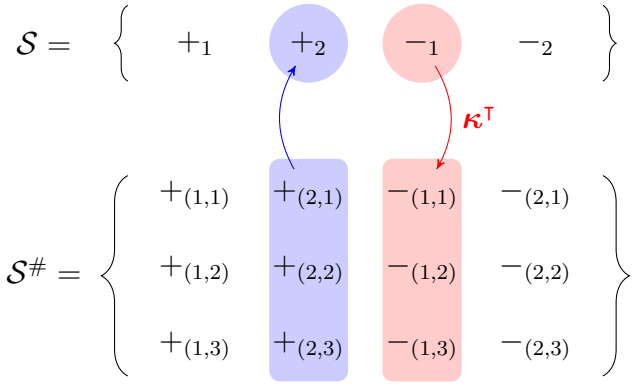


Figure 5.1: A visualisation of the behaviour of $\widetilde{\varphi}^\#(t)$ at level ζ for $\ell = 3$. States marked with $+/-$ are upwards/downwards phases, respectively. When the level hits ζ from above, $\widetilde{\varphi}^\#(t)$ transitions from $\mathcal{S}_-$ to $\mathcal{S}_-^\#$ according to initial distribution vector $\boldsymbol{\kappa}^\intercal$; when the level hits ζ from below, $\widetilde{\varphi}^\#(t)$ transitions from $\mathcal{S}_+^\#$ to $\mathcal{S}_+$, retaining the phase information, but losing the state of the horizon process.

5.2.1 Time-dependent quantities in the Parisian model

We can now define bounded variants of the time-dependent quantities $\widehat{\boldsymbol{\Psi}}^\#(s)$, $\widehat{\boldsymbol{\Xi}}^\#(s)$, $\widehat{\mathbf{G}}^\#(s;x)$ and $\widehat{\mathbf{H}}^\#(s;x)$, but we are only interested in an analogue of ${}^z_0\widetilde{\boldsymbol{\Xi}}(s)$, which is the LST of the time taken to return to the level z in the bounded model $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$, starting from a downwards phase.

Mirroring Theorem 3 from Bean *et al.* [13], we construct the quantity ${}^z_0\widetilde{\boldsymbol{\Xi}}^\#(s)$, which is the LST of the time taken to return to the level z in the Parisian model $\{\widetilde{\mathcal{F}}^\#(t)\}_{t\geq 0}$. Here we assume that $z \leq \zeta$, so we need not worry about the level crossing ζ—later, we set $z = \zeta$ anyway.

Define $\widetilde{\theta}^\#(x) = \inf\{t > 0 : \widetilde{M}^\#(t) = x\}$. Then for $\boldsymbol{j} \in \mathcal{S}_-^\#$, $\boldsymbol{i} \in \mathcal{S}_+^\#$:

$$\left[{}^z_0\widetilde{\boldsymbol{\Xi}}^\#(s)\right]_{\boldsymbol{j},\boldsymbol{i}} = \mathbb{E}\left[e^{-s\widetilde{\theta}^\#(z)}\mathbb{I}\left(\widetilde{\theta}^\#(z) < H, \widetilde{\varphi}^\#(\widetilde{\theta}^\#(z)) = \boldsymbol{i}\right) \,\middle|\, \widetilde{M}^\#(0) = z, \widetilde{\varphi}^\#(0) = \boldsymbol{j}\right].$$

Theorem 3 from Bean *et al.* [13] implies that

$${}^z_0\widetilde{\boldsymbol{\Xi}}^\#(s) = {}^z_0\widehat{\mathbf{H}}^\#_{-+}(s;z) + {}^z_0\widehat{\mathbf{G}}^\#_{--}(s;z){}^z_0\breve{\mathbf{W}}^\#(s){}^z_0\widehat{\mathbf{H}}^\#_{++}(s;0), \tag{5.3}$$

where

$${}^z_0\breve{\mathbf{W}}^\#(s) = \mathbf{V}^\#(s)\left(\mathbf{I} - {}^z_0\widehat{\mathbf{G}}^\#_{+-}(s;0)\mathbf{V}^\#(s)\right)^{-1},$$

and

$$\mathbf{V}^\#(s) = \breve{\mathbf{P}}^\#_{-0}(s\mathbf{I} - \breve{\mathbf{T}}^\#_{00})^{-1}\breve{\mathbf{T}}^\#_{0+} + \breve{\mathbf{P}}^\#_{-+}.$$

The quantities $\mathbf{V}^{\#}(s)$ and ${}^{z}_{0}\widetilde{\mathbf{W}}^{\#}(s)$ have the exact same interpretation as their analogues in Section 2.4.2: $\mathbf{V}^{\#}(s)$ is the LST of the time spent during a visit to the lower boundary of 0, and ${}^{z}_{0}\widetilde{\mathbf{W}}^{\#}(s)$ is the LST of the time spent during a path which traverses the interval $[0, z)$, beginning and ending at level 0, and returning to level 0 any number of times, as long as the path never hits the upper threshold z.

Again, the departure from Bean *et al.* [13] is the inclusion of the horizon, which can evolve when the level is in the interval $(0, z)$, and at the lower boundary.

5.2.2 The Parisian time to mortality

We now attempt to derive an expression for the TTM under the Parisian model, using Lemma 5.3 and Lemma 5.4 as stepping stones.

Lemma 5.3. *Suppose the fluid process $\{\widetilde{\mathcal{F}}^{\#}(t)\}_{t\geq 0}$ hits level ζ from above in phase $i \in \mathcal{S}_-$. Then the LST of the time to return to ζ before the horizon is reached, in phase $j \in \mathcal{S}_+$, is $[\widetilde{\mathbf{D}}(s)]_{i,j}$, where*

$$\begin{aligned}\widetilde{\mathbf{D}}(s) &= \check{\mathbf{P}}_{\zeta}\, {}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}^{\#}(s)\, \widehat{\mathbf{P}}_{\zeta} \\ &= (\mathbf{I}_- \otimes \boldsymbol{\kappa}^{\intercal})\, {}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}^{\#}(s)\, (\mathbf{I}_+ \otimes \mathbf{1}_{\ell})\,.\end{aligned} \tag{5.4}$$

Proof. Consider a path that has just hit level ζ in phase $i \in \mathcal{S}_-$ and will end at level ζ in phase $j \in \mathcal{S}_+$ at some time prior to H:

- At the moment the level hits ζ from above in phase $i \in \mathcal{S}_-$, the phase transitions to some phase $\boldsymbol{l} \in \mathcal{S}_-^{\#}$ according to $[\check{\mathbf{P}}_{\zeta}]_{i,\boldsymbol{l}}$.
- From level ζ in phase $\boldsymbol{l}$, the process undertakes a journey below ζ, and must return to ζ in some phase $\boldsymbol{l}' \in \mathcal{S}_+^{\#}$ before the horizon is reached. The LST of the time taken on this journey is $[{}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}^{\#}(s)]_{\boldsymbol{l},\boldsymbol{l}'}$.
- At the moment the level hits ζ from below in phase $\boldsymbol{l}'$, the phase transitions to $j \in \mathcal{S}_+$ according to $[\widehat{\mathbf{P}}_{\zeta}]_{\boldsymbol{l}',j}$.

Thus, the LST of the time taken to return to ζ before the horizon is reached, starting from $i \in \mathcal{S}_-$ and ending in $j \in \mathcal{S}_+$ is

$$\begin{aligned}[\widetilde{\mathbf{D}}(s)]_{i,j} &:= \sum_{\boldsymbol{l}\in\mathcal{S}_-^{\#}} \sum_{\boldsymbol{l}'\in\mathcal{S}_+^{\#}} [\check{\mathbf{P}}_{\zeta}]_{i,\boldsymbol{l}}\, [{}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}^{\#}(s)]_{\boldsymbol{l},\boldsymbol{l}'}\, [\widehat{\mathbf{P}}_{\zeta}]_{\boldsymbol{l}',j} \\ &= \left[\check{\mathbf{P}}_{\zeta}\, {}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}^{\#}(s)\, \widehat{\mathbf{P}}_{\zeta}\right]_{i,j}.\end{aligned}$$

■

Lemma 5.4. *The LST of the time until the horizon is reached, given that the process starts at level ζ in a downwards phase and mortality occurs, is*

$$\boldsymbol{\Omega}(s) = \int_{t=0}^{\infty} e^{-st} \left(\int_{r=t}^{\infty} \boldsymbol{\xi}(r)\, dr\, \mathbf{1}_+ \right) \boldsymbol{\kappa}^\intercal e^{\mathbf{K}t} \boldsymbol{k}\, dt,$$

where $\boldsymbol{\xi}(r)$ is the (matrix) probability density function of returning to level ζ starting from a downwards phase, obtained by inverting the LST ${}^{\zeta}_{0}\widetilde{\boldsymbol{\Xi}}(s)$.

Proof. The LST $\boldsymbol{\Omega}(s)$ is the final component of the TTM, corresponding to the final journey below ζ which leads to mortality (i.e., the horizon is reached before the process returns to ζ). We obtain $\boldsymbol{\Omega}(s)$ by taking the LST with respect to the joint distribution $\boldsymbol{F}(t,r) = \mathbb{P}\Big[H \leq t,\, \widetilde{\theta}(\zeta) \leq r\Big]$, for $t \leq r$, thus imposing that mortality must occur. In a slight abuse of notation, let $\boldsymbol{F}(dt, dr)$ denote the corresponding joint Riemann-Stieltjes integrator, and so

$$\begin{aligned}\boldsymbol{\Omega}(s) &= \int_{t=0}^{\infty} \int_{r=t}^{\infty} e^{-st}\, \boldsymbol{F}(dt, dr) \\ &= \int_{t=0}^{\infty} e^{-st} \int_{r=t}^{\infty} d\mathbb{P}\Big[\widetilde{\theta}(\zeta) \leq r\Big]\, d\mathbb{P}[H \leq t]\,,\end{aligned}$$

due to the independence of H and $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$. Further,

$$\boldsymbol{\Omega}(s) = \int_{t=0}^{\infty} e^{-st} \left(\int_{r=t}^{\infty} \boldsymbol{\xi}(r)\, dr\, \mathbf{1}_+ \right) \boldsymbol{\kappa}^\intercal e^{\mathbf{K}t} \boldsymbol{k}\, dt,$$

since $H \sim \text{PH}(\boldsymbol{\kappa}^\intercal, \mathbf{K})$. We assume that $\mathbb{P}[\widetilde{\theta}(\zeta) \leq r]$ is continuous[1] and replace $d\mathbb{P}[\widetilde{\theta}(\zeta) \leq r]$ with $\boldsymbol{\xi}(r)dr$, and hence, obtain the result.

Note that we are not interested in what phase the process ends up in at the time of mortality, and so we account for this by post-multiplying the inner integral by $\mathbf{1}_+$. The PDF of the horizon is a scalar function, and already accounts for this. ■

Theorem 5.5. *The LST of the TTM under the Parisian model with ℓ horizon stages, given that the process starts in phase $i \in \mathcal{S}_-$ at level b, is*

$$\widehat{F}^{\ell}_{\text{TTM}}(s) = \boldsymbol{e}_i^\intercal\, {}^{b}_{\zeta}\widetilde{\mathbf{G}}_{--}(s; b) \left(\mathbf{I} - \widetilde{\mathbf{D}}(s)\, {}^{b}_{\zeta}\widetilde{\boldsymbol{\Psi}}(s) \right)^{-1} \boldsymbol{\Omega}(s), \tag{5.5}$$

where $\widetilde{\mathbf{D}}(s)$ and $\boldsymbol{\Omega}(s)$ are as defined in Lemma 5.3 and Lemma 5.4 respectively.

[1] If $\mathbb{P}[\widetilde{\theta}(\zeta) \leq r]$ has discontinuities, then the expression for $\boldsymbol{\Omega}(s)$ will contain additional terms corresponding to the point masses of $\mathbb{P}[\widetilde{\theta}(\zeta) \leq r]$ at these discontinuities.

Proof. Consider a path which begins at level b in phase $i \in \mathcal{S}_-$, and eventually reaches the cemetery state 0, at which point mortality occurs:

- For mortality to occur, the level must drain from b to ζ, possibly returning to level b before reaching ζ. The LST of the time taken for this journey is ${}^{b}_{\zeta}\widetilde{\mathbf{G}}_{--}(s;b)$. Since the initial phase is $i \in \mathcal{S}_-$, we pre-multiply by the (row) unit vector $\boldsymbol{e}_i^\intercal$, which only takes the i^{th} row of this matrix.

- Once the level hits ζ, the process can make any number of journeys below ζ, followed by a journey above ζ, as long as the journey below ζ returns to ζ before the horizon is reached. The LST of the journey below ζ is $\widetilde{\mathbf{D}}(s)$ from Lemma 5.3, and the LST of the journey above ζ is ${}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)$. Thus, the LST of the full journey is $\widetilde{\mathbf{D}}(s){}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)$ by the convolution theorem (since the times for each segment are conditionally independent). Therefore, the LST of any number of such journeys is $\sum_{m=0}^{\infty}[\widetilde{\mathbf{D}}(s){}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)]^m = \left(\mathbf{I} - \widetilde{\mathbf{D}}(s){}^{b}_{\zeta}\widetilde{\mathbf{\Psi}}(s)\right)^{-1}$.

- The process must eventually make a single journey below ζ which ends in mortality. The LST of the time spent on this journey before the instant of mortality is $\mathbf{\Omega}(s)$ from Lemma 5.4.

By conditional independence and the convolution theorem, the result follows. ■

This construction is almost identical to the construction in Section 3.3.2, which considers a single path eventually leading to mortality. The only difference is that the time that the process is allowed to spend below the critical algal threshold before mortality is now random, whereas it was previously constant.

Interpreting and computing $\mathbf{\Omega}(s)$

The LST $\mathbf{\Omega}(s)$ is more complicated than the corresponding component of the TTM in Section 3.3.2, but still has an intuitive explanation.

The outer integral takes the LST of the time until the horizon is reached: if there was no condition on the level process, this would be the LST of a PH$(\boldsymbol{\kappa}^\intercal, \mathbf{K})$ distribution; however, we need to account for the fact that the level might return to ζ before the horizon expires. This is done by the inner integral: we multiply by the the probability that, for a given value of t, the level process has not yet returned to level ζ. This is the the complementary cumulative distribution function (CCDF) of the time until the level returns to ζ, which we denote $\Xi(t) = \int_{r=t}^{\infty} \boldsymbol{\xi}(r)\, dr$.

Since we only consider the process $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$ here, the LST $\mathbf{\Omega}(s)$ has dimension $s_- \times 1$. Thus, unlike the quantity $\widetilde{\mathbf{D}}(s)$ we do not need an extra matrix to account

for the transition from $\mathcal{S}_-$ to $\mathcal{S}_- \times \mathcal{S}_\ell$ in Theorem 5.5. Note that the PDF of the horizon is a scalar quantity: all of the phase behaviour is accounted for by the inner integral, which need only account for the initial downwards phase.

The most difficult part of $\mathbf{\Omega}(s)$ to compute is the outer integral in Lemma 5.4, which requires us to truncate the integral at a finite value and approximate the integral somehow. This can be done by using trapezoidal integration over some interval $[0, T_{\max}]$, where $T_{\max}$ is sufficiently large, using a sufficient number of (optionally, equally spaced) points in $[0, T_{\max}]$. We face some major issues, however.

Computing $\mathbf{\Omega}(s)$ requires evaluating the CCDF $\Xi(t)$ at every point in our interval of choice, which boils down to inverting the LST ${}_0^{\zeta}\widetilde{\mathbf{\Xi}}(s)$ at each of these points. If we require a high degree of accuracy, or a large value of $T_{\max}$ to capture the horizon process, then this computation may become very taxing.

Moreover, computing $\widetilde{\mathbf{D}}(s)$ involves solving Sylvester equations with matrices which scale in size with ℓ. Even for $\ell \approx 10$, these computations take their toll, and it may even become intractable to invert the LST of the TTM for much larger ℓ.

For the above reasons, we decide that LST inversion may not be the most effective way to obtain the distribution of the TTM in this case, particularly since we would like to test the effects of ℓ. Rather, we consider simulation as an alternative. The main drawback of simulation is that simulating a model with a low probability of mortality can be slow—although not prohibitively so. Variance reduction techniques such as importance sampling can speed up simulation, but we leave this as an avenue for future research. Overall, simulation is an effective means to obtain the distribution of the TTM, and is the focus of the next section.

5.3 Erlangization

We now have a model to simulate, but before we can do any simulation we need a horizon distribution. The Parisian framework allows for any PH distribution—of which there are an infinite number—so which one do we choose?

Closer to the start of this chapter, we learned a bit about the history of Parisian ruin, largely born in Asmussen, Avram and Usabel [4], which pioneered the idea of adding an independent horizon to a fluid model. The paper focused on the case where the horizon had an Erlang distribution, which greatly simplified computations of the probability of (Parisian) ruin. This method was later coined *Erlangization* in Stanford *et al.* [70], which extended the method to obtain the distribution of the deficit of the risk process when ruin occurs.

Erlangization is a nice way to demonstrate the capabilities of our Parisian model. The model is a simple extension to the basic model from Chapter 3, since

The Erlang distribution

The Erlang distribution is widely used in probability—particularly in the modelling of telecommunications networks. For a set of ℓ independent exponential random variables with common rate μ (and hence a mean of $1/\mu$), the sum of these random variables has an Erlang(ℓ, μ) distribution with mean ℓ/μ.

Important: here μ denotes the *rate* parameter, not the scale parameter.

In Figure 5.2, this is illustrated as a continuous-time Markov chain (CTMC), starting in state ℓ: transitions between states in a CTMC are exponentially distributed, and thus the sum of these transition times in the illustrated CTMC is Erlang distributed. This sum is the time of absorption into state 0, which is precisely the definition of a phase-type (PH) distribution.

The Erlang distribution is the PH distribution with the lowest variance for a given number of phases [26]. As a consequence of the Central Limit Theorem, as $\ell \to \infty$, with $\mu = \ell/\tau$ the distribution converges to a normal distribution centred around τ, with variance τ^2/ℓ, which itself approximates the Dirac delta distribution centred at τ.

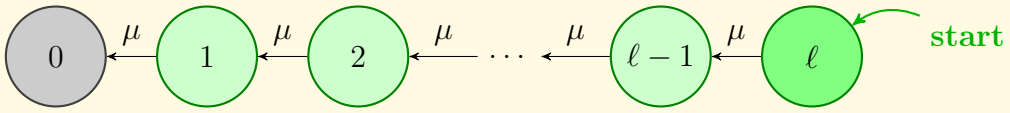


Figure 5.2: An Erlang(ℓ, μ) distributed random variable has the same distribution as the absorption time of the above CTMC, starting in state ℓ. If $\mu = \ell/\tau$, then the mean of this distribution is τ for any $\ell \in \mathbb{N}$.

it only requires two additional parameters to describe the horizon: ℓ, the number of Erlang stages, and μ, the *rate* of the distribution, whence ℓ/μ is the mean.

A brief introduction to the Erlang distribution is shown in the above insert The Erlang distribution, which illustrates the distribution as a CTMC.

Furthermore, Asmussen, Avram and Usabel [4] showed that the probability of ruin in an Erlangized fluid model converges to the probability of fixed time ruin as the number of Erlang stages increases to infinity. Fixed time ruin is precisely what we studied in Chapter 3 and Chapter 4, therefore, we expect that as $\ell \to \infty$, the distribution of the TTM in an Erlangized fluid model with $\mu = \ell/\tau$ converges to the distribution of the original TTM with parameter τ.

In the remainder of this chapter, we construct an Erlangized fluid model and simulate it to demonstrate this convergence; however, moving from ruin probabilities to the distribution of time until ruin yields some additional complexities.

5.3.1 Model specification and simulation setup

We will be comparing the distributions of the TTM in two models:

- Model A: the basic fluid model, introduced in Chapter 3, using the canonical parameters from Section 3.2.4.
- Model B: the Erlangized fluid model with the same fluid model parameters as Model A, but with an additional horizon specification $H \sim \mathrm{PH}_\ell(\boldsymbol{\kappa}^\intercal, \mathbf{K})$.

To create an *Erlangized* fluid model, H must be an Erlang random variable. We achieve this with the parameters

$$\boldsymbol{\kappa}^\intercal = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}, \tag{5.6}$$

$$\mathbf{K} = \begin{bmatrix} -\mu & 0 & 0 & 0 & \dots & 0 & 0 \\ \mu & -\mu & 0 & 0 & \dots & 0 & 0 \\ 0 & \mu & -\mu & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\mu & 0 \\ 0 & 0 & 0 & 0 & \dots & \mu & -\mu \end{bmatrix}, \tag{5.7}$$

$$\mu = \frac{\ell}{\tau}.$$

The row vector $\boldsymbol{\kappa}^\intercal$ has $\ell \in \mathbb{N}$ elements, $\mathbf{K}$ is a $\ell \times \ell$ matrix and $\tau > 0$.

To show that Model B converges to Model A as $\ell \to \infty$, we need to think about what *convergence* means in this context. Here, we say that convergence occurs if the cumulative distribution function (CDF) of the TTM under Model B converges to the CDF under Model A.

To obtain the CDF of the TTM for Model A, we perform LST inversion using Algorithm B.3. Here, we use the CME method without shifting ($n = 50$ nodes). To obtain the CDF of the TTM for Model B, we simulate the model with a modified version of our existing simulation code: we create the augmented model and add a way to check if the cemetery state has been reached.

We can simulate the model with a modified version of Algorithm B.4. This new function will be almost identical, but with some minor changes: a new generator $\mathbf{T}'$ is created at (†), and we check if the process has entered the cemetery state at (‡). The matrix $\mathbf{T}'$ is defined as

$$\mathbf{T}' = \begin{bmatrix} 0 & \mathbf{0}_{1\times\ell} \\ -\mathbf{T}^{\#}\mathbf{1}_\ell & \mathbf{T}^{\#} \end{bmatrix}.$$

Creating the Parisian model ($\mathbf{T}^{\#}$, $\mathbf{C}^{\#}$, etc.) is straightforward, and can be achieved by following the construction in Section 5.2. This construction is performed by Algorithm B.5. Model B is created from the Parisian model, with the addition of $\boldsymbol{\kappa}^{\intercal}$ and $\mathbf{K}$ as in Eq. (5.6) and Eq. (5.7), respectively.

5.3.2 Convergence to the original model

We proceed to obtain the CDF of the TTM for both models, using 50 nodes in the LST inversion (Model A) and 20,000 simulations for each value of ℓ (Model B).

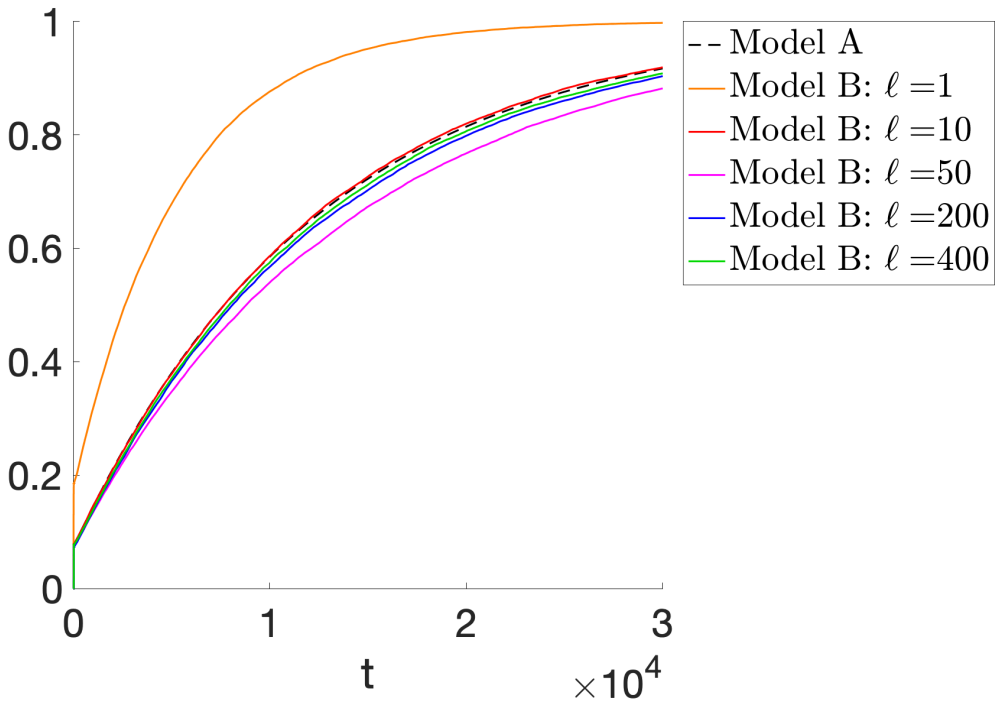
Figure 5.3 shows the CDF of the TTM for both models. For Model B, we use values of $\ell = 1, 10, 50, 200$ and 400. For higher values of ℓ, the CDF appears to converge to the basic model as expected; however, it is worth pointing out that when $\ell = 10$, the CDF appears to be much more accurate than it should be for so few Erlang stages, and that the convergence of the CDFs is non-monotonic. The explanation for this anomaly lies in a shorter time window.

Figure 5.3b shows the CDFs for both models for the same parameters, but on a much shorter time scale. We observe some interesting behaviour around $t = 15$.

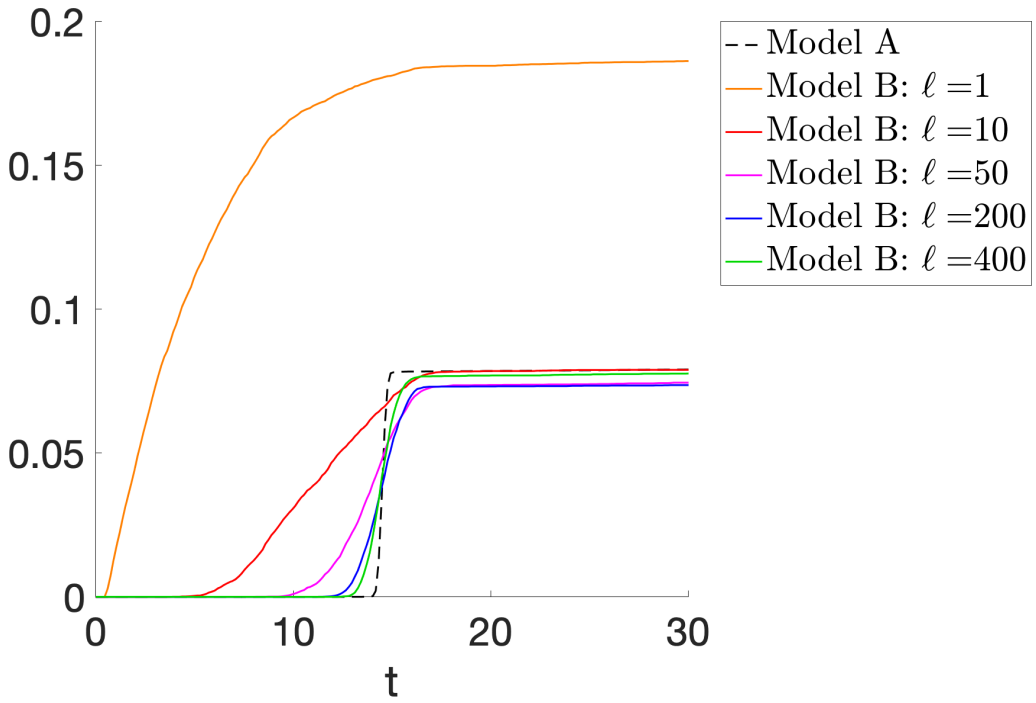
Recall from Section 4.4.1 that in Model A, mortality can occur at a minimum time $t_{\text{crit}} = 14.5$, and that the CDF of the TTM has a discontinuity at t_{crit}. In Figure 5.3b, we observe this discontinuity in the CDF for Model A, but not in the CDFs for Model B. This is because in Model B, mortality can occur before τ, since mortality is no longer fixed, and thus mortality can occur before t_{crit}. However, as ℓ increases, we observe the CDFs for Model B beginning to approximate the discontinuity at t_{crit}.

This is exactly as expected, due to the fact that an Erlang distribution converges to a normal distribution centred around τ, with variance τ^2/ℓ. As ℓ increases, the Erlang distribution should converge to a Dirac delta distribution centred at τ, which approximates a constant random variable, taking the value τ. Thus, as ℓ increases, the minimum time to mortality in the Erlangized model should approach t_{crit}, since the distribution of time to reach ζ from level 1 is exactly the same in Models A and B.

When $\ell = 1$, the CDF is clearly a poor approximation, but as ℓ increases, we observe the CDF for Model B becoming steeper at t_{crit}, and containing less probability density prior to t_{crit}. An artefact of this change is that when $\ell = 10$, the CDF for Model B just happens to provide a very good approximation to the CDF for Model A after t_{crit}, despite doing a poor job of approximating the discontinuity. As a result, the convergence of Model B to Model A appears to be monotonic, in a certain sense, when viewed on a smaller time scale.



(a) Long time scale. The CDF for Model B appears to converge as ℓ increases, but the CDF also appears to be surprisingly accurate when $\ell = 10$.



(b) Short time scale. The behaviour of the CDFs at $t_{\text{crit}} = 14.5$ is particularly interesting: as ℓ increases, the CDF for Model B better approximates the discontinuity at t_{crit}.

Figure 5.3: CDF of the TTM for Model A, obtained using the CME method with $n = 50$ nodes, and Model B, obtained with 20,000 simulations for each CDF, for $\ell = 1, 10, 50, 200$ and 400.

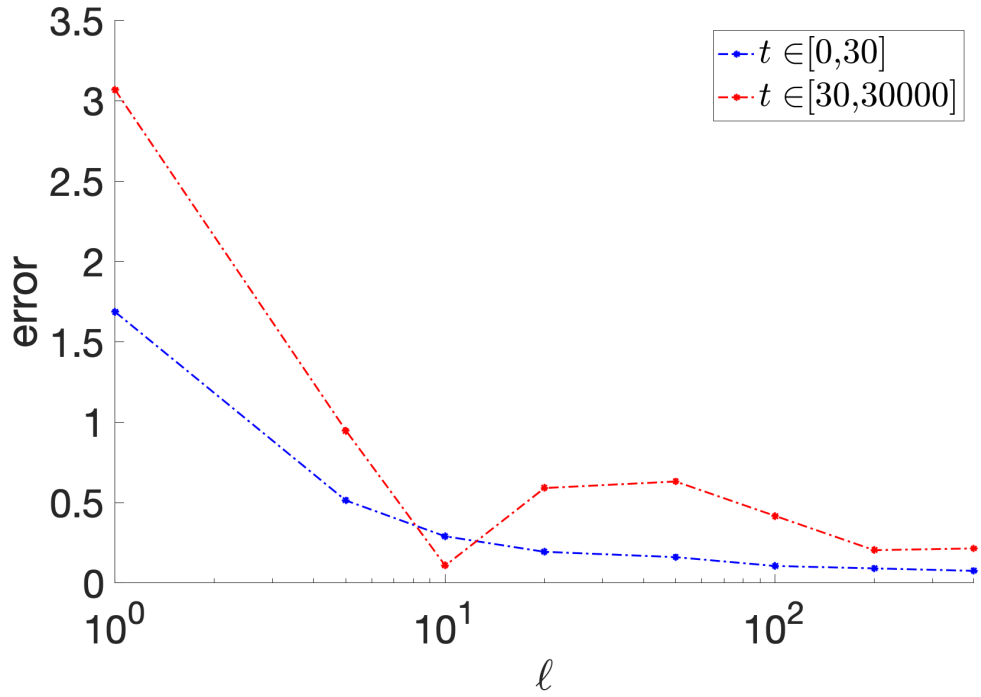


Figure 5.4: Error between Model A and Model B on two time scales—short and long—calculated as the ℓ^2-norm between the discretised CDFs for each model, $\left\| \boldsymbol{f}_A - \boldsymbol{f}_B^{(\ell)} \right\|_2$. Model B CDFs are obtained using 20,000 simulations each. Error is evaluated at $\ell = 1, 5, 10, 20, 50, 100, 200$ and 400, shown by asterisks; dot-dashed lines are shown to help illustrate trends. Horizontal axis is a logarithmic scale.

Error as a function of ℓ

Denote the CDFs of Model A and Model B, for ℓ Erlang stages, by $F^A(t)$ and $F_\ell^B(t)$, respectively. Similarly to the approach in Section 4.4.2, we have discretised $F^A(t)$ and $F_\ell^B(t)$ to obtain the sequences $\boldsymbol{f}_A$ and $\boldsymbol{f}_B^{(\ell)}$, respectively. We then take the ℓ^2-norm[2] of these sequences, $\left\| \boldsymbol{f}_A - \boldsymbol{f}_B^{(\ell)} \right\|_2$, which is a proxy for the error between the two CDFs. We choose to evaluate this error on two sets: $T_1 = [0, 30]$ (short time scale) and $T_2 = [30, 30000]$ (long time scale).

Figure 5.4 shows the error as a function of ℓ (on a logarithmic scale), using 20,000 simulations for each CDF in Model B and 200 equally spaced points to construct the sequences $\boldsymbol{f}_A$ and $\boldsymbol{f}_B^{(\ell)}$. On a long time scale, we see that the error is very small at $\ell = 10$, in fact smaller than at $\ell = 400$, coinciding with our previous findings. On a shorter time scale, the error appears to be steadily decreasing in ℓ, indicating that the Erlangization approach is approximating the jump at t_{crit} better as ℓ increases, also in agreement with our previous observations.

[2]The ℓ^2-norm is defined in Section 4.4.2. The 'ℓ' symbol in 'ℓ^2-norm' is unrelated to the parameter ℓ used in this chapter.

5.3.3 Discussion, conclusions and the future

The CDF of the TTM under Model B certainly appears to converge to the distribution of the TTM under Model A, confirming that the distribution of time until Parisian ruin converges, in some sense, to the distribution of time until fixed-time ruin. The behaviour of this convergence offers more complex behaviour than might be expected, both due to the difficulties in comparing distributions, rather than sequences of numbers, and the need to encapsulate the dynamics of the distribution across multiple time scales. When looking at the distributions on a small scale, the convergence appears to be monotonic, but this is not the case when looking at the distributions on a longer scale.

Recall from Section 4.4.4 that on a short time scale, the CDF tells us about the probability of a coral colony dying almost immediately due to the initial bleaching event at $t = 0$, or the proportion of colonies that do so. On a long time scale, the CDF tells us about the probability of a colony surviving up until that time, or the proportion of colonies to survive that long.

With this in mind, it is reassuring that Model B tends toward Model A on both time scales, as ℓ increases, since both interpretations are important.

Improving upon the Parisian model

The Parisian model, under which Erlangization provides a nice means to analyse, is a versatile model which we can use to analyse the system of coral-algal symbiosis. The ability to model a random time of mortality enhances the existing model for mortality which we explored in Chapters 3 and 4; however, there is still much room for improvement.

At the end of Chapter 3, we mentioned the lack of consideration for the energy process of the coral host, which has still not been accounted for in the Parisian model. Further, we would like to answer questions about the survivability of the coral host which take symbiont diversity into account, but our model does not yet account for the crucial detail that genus C is better at providing nutrients to the coral host than genus D. This is what we will explore in the next chapter, armed with a newly acquired arsenal of horizon processes.

Chapter 6

An Energy Model

In this chapter, we introduce a more sophisticated model for coral-algal symbiosis, building on the model from Chapter 3, and incorporating features of the model from Chapter 5. We aim for this model to be more biologically authentic, so that we are able to better incorporate the dynamics between different symbiont genera. This provides an avenue to answer the research questions posed in Section 2.1.5. We first discuss a new multi-threshold methodology enabling us to further distinguish between the qualities of each symbiont genus. We then proceed to shape this concept into a model by extending the capabilities of the Parisian ruin model from Chapter 5 to incorporate a two-way horizon, and add level dependent behaviour into the framework. We define the LST of the TTM, and once again use simulation to obtain results in our model. We analyse the new model in detail, using tail probabilities to compare the model to single-genus models of a similar construction. We conclude with a discussion and interpretation of these results in the context of coral bleaching on the GBR, and of our modelling processes.

6.1 A new methodology

Recall, from the end of Chapter 3, that the energy process $E(t)$ of our fluid model defines the critical threshold ζ at which the total rate of energy production is equal to the total rate of energy consumption in the coral host. Specifically, if $E'(t) = \widetilde{M}(t)P_{\widetilde{\varphi}(t)} - R$ (where $P_{\widetilde{\varphi}(t)}$ is the energy production rate in phase $\widetilde{\varphi}(t)$ and R is the energy requirement rate of the coral host), then ζ is the level such that $\zeta P_{\widetilde{\varphi}(t)} = R$. It is implicitly assumed that $P_{\widetilde{\varphi}(t)}$ is the same for all phases. However, genus C has a higher photochemical efficiency than genus D [22, 25]: in our modelling framework, this means that $P_{\widetilde{\varphi}(t)}$ is *not* the same for all phases.

We are interested in answering questions about dynamics between different symbiont genera: to do so, it is crucial to capture this feature in our model.

One way to do this is by specifying an energy production rate for each genus: P_C for genus C and P_D for genus D. In our model, this means that for any phase where C is the dominant genus, the energy production rate is P_C, and similarly for genus D. Subsequently, we may define two critical thresholds ζ_C and ζ_D such that $\zeta_C P_C = \zeta_D P_D = R$, with a similar interpretation to the threshold ζ from the basic model: when the level is below ζ_C and genus C is the dominant genus, or when the level is below ζ_D and genus D is the dominant genus, the coral host will be in energy deficit; otherwise, the coral host will have an energy surplus.

Note that the fact that genus C has a higher energy production rate than genus D implies $0 < \zeta_C < \zeta_D < 1$.

This new structure also requires a new definition of mortality, however, it is not immediately obvious what definition to choose. We could decree that under ζ_D, only the time spent in genus D counts towards mortality, and that transitioning into genus C 'resets the clock'. This is a perfectly good definition, and consistent with that of the singular-threshold model, however, it means that this 'resetting' can occur at an *arbitrary level*, as opposed to a fixed level ζ.

There are several types of sample paths in which this can occur, for example: a path where genus D is the dominant genus, the level process hits level ζ_D from above and then genus C becomes the dominant genus; alternatively, where genus C is the dominant genus, the level dips below ζ_D and then genus D becomes the dominant genus.

What happens to the mortality process in these examples?

Moreover, if genus D is dominant and the level is in between ζ_C and ζ_D, the clock is reset when the process transitions to genus C, and the process has potentially forgotten about the large amount of time spent in genus D, during which the coral host is in energy deficit.

In the single-threshold model, it is reasonable to assume this 'memoryless' property, but in a multi-threshold model, it is less reasonable. With two thresholds, the coral can replenish its energy by a simple change in genus, as opposed to returning to the critical threshold.

Ideally, we would like to give the process some sort of 'memory', so that in cases like the ones above, the previous bleaching history has an effect on the future of the coral's health. Even more ideally, we would like to create an *explicit energy process*, rather than approximating this process by resetting the mortality clock at the critical threshold. In this chapter, both of these features shall be incorporated.

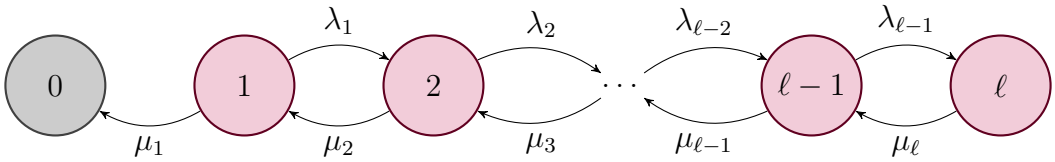


Figure 6.1: A visualisation of the two-way horizon H as a CTMC on the state space $\mathcal{S}_\ell \cup \{0\}$. This is similar to the Erlang horizon in Figure 5.2.

6.2 The energy model

In this section we introduce a model which captures the same behaviour as the basic model, with the addition of an *explicit energy process*. Moreover, this model still lies within the stochastic fluid framework, and thus much of the work from the basic model can be translated over into the new environment.

We refer to this model as the *energy model*.

The energy model combines a level-dependent bounded fluid model with a phase-type (PH) horizon. We define the latter in this section, and then add level-dependent behaviour in Section 6.3.

6.2.1 A two-way horizon

Define a horizon $H \sim \text{PH}(\boldsymbol{\kappa}^\intercal, \mathbf{K})$ on the state space $\mathcal{S}_\ell = \{1, 2, \ldots, \ell\}$ with absorbing state 0, for some fixed $\ell \in \mathbb{Z}_+$. Suppose $\mathbf{K}$ has the following structure:

$$\mathbf{K} = \begin{bmatrix} -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \ldots & 0 & 0 \\ \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \ldots & 0 & 0 \\ 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & -(\lambda_{\ell-1}+\mu_{\ell-1}) & \lambda_{\ell-1} \\ 0 & 0 & 0 & 0 & \ldots & \mu_\ell & -\mu_\ell \end{bmatrix},$$

where $\lambda_i, \mu_i \in (0, \infty)$, $\forall i \in \mathcal{S}_\ell$.

The horizon H can then be interpreted as the absorbing time of a CTMC on the state space $\mathcal{S}_\ell \cup \{0\}$, which is a birth-death process with absorbing state 0, as illustrated in Figure 6.1. The process has infinitesimal generator

$$\mathbf{K}' = \begin{bmatrix} 0 & \mathbf{0}_{1\times\ell} \\ \mathbf{k} & \mathbf{K} \end{bmatrix}$$

where $\mathbf{k} = -\mathbf{K}\mathbf{1}_\ell$ is the vector of absorbing rates ($\mathbf{k}$ will be all zeros except for a single non-zero element μ_1, since the process can only be absorbed from state 1).

In Section 5.1, we explored the idea of Parisian ruin, which fit snugly into our modelling framework and allows for a more general definition of the TTM. In the Parisian model, mortality occurs when the horizon is reached; moving towards the cemetery state represents getting closer to death. This 'progress' resets upon returning to the threshold. Cumulative Parisian ruin would be the other extreme: the 'progress' towards mortality persists across visits to the threshold.

As of yet, we have only thought of this progress towards mortality as an abstract means to capture the entire mortality process, but in the context of the coral model, this state has an interpretation: *the stored energy of the coral.*

By incorporating the energy process into the phase process, we can discretise the energy buffer $E(t)$. If the maximum energy storage of the coral host is $E_{\max}$, then when the energy stage is j, the amount of energy stored is $jE_{\max}/\ell$. We can increase ℓ to better approximate a (more realistic) continuous energy buffer.

6.2.2 Another augmented model

We now define a fluid model using the two-way horizon. This model is similar to the model in Chapter 5, but the horizon is now slightly more sophisticated, and can be directly interpreted as the energy buffer of the coral host.

Consider the fluid model $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0} = \{(\widetilde{M}(t), \widetilde{\varphi}(t)) : t \geq 0\}$. Here, $\widetilde{M}(t)$ denotes the level process, bounded on the interval $[0, b]$ and $\widetilde{\varphi}(t)$ denotes the phase process on the state space $\mathcal{S}^{\mathcal{E}} \cup \{0\} = (\mathcal{S} \times \mathcal{S}_\ell) \cup \{0\}$. We order $\mathcal{S}^{\mathcal{E}}$ lexicographically: $(i_1, i_2) \leq (j_1, j_2) \iff i_1 < j_1$ or $i_1 = j_1, i_2 \leq j_2$.

When $\widetilde{\varphi}(t) = (i, j) \in \mathcal{S}^{\mathcal{E}}$, we refer to $j \in \mathcal{S}_\ell$ as the *energy stage.*

The phase $\varphi(t) = (i_1, i_2) \in \mathcal{S}^{\mathcal{E}}$ can be interpreted as follows: i_1 represents the dominant symbiont genus, and whether growth or bleaching is occurring; the energy stage i_2 represents the amount of energy stored by the coral host (manifested in reality as the volume of lipids), where ℓ represents the maximum amount of stored energy. Mortality occurs when the phase process reaches state 0, representing the coral's energy being fully depleted.

The behaviour of this model is very similar to that of the bounded fluid model in Section 2.4.2. The idea here is to add a horizon to an existing bounded fluid process $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$, and thus we can assume that $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$ has already been defined by matrices $\mathbf{C}$, $\widehat{\mathbf{P}}$, $\check{\mathbf{P}}$, $\mathbf{T}$, $\widehat{\mathbf{T}}$ and $\check{\mathbf{T}}$. We define the fluid input/output rate matrix for the augmented process as $\mathbf{C}^{\mathcal{E}} = \mathbf{C} \otimes \mathbf{I}_\ell$, and so the level process ignores the energy stage. Therefore, the process $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0}$ evolves as follows:

- When $\widetilde{M}(t) \in (0, b)$ and $\widetilde{\varphi}(t) = (i_1, i_2) \in \mathcal{S}^{\mathcal{E}}$, the level changes at rate c_{i_1}, and the phase process evolves according to $\mathbf{T}^{\mathcal{E}}$.

- When the level $\widetilde{M}(t)$ hits the lower boundary 0, the phase process immediately enters $(\check{\mathcal{S}}_0 \times \mathcal{S}_\ell) \cup (\mathcal{S}_+ \times \mathcal{S}_\ell)$ according to the matrix

$$\check{\mathbf{P}}^{\mathcal{E}} = \left[\check{\mathbf{P}}_{-0} \otimes \mathbf{I}_\ell \;\middle\vdots\; \check{\mathbf{P}}_{-+} \otimes \mathbf{I}_\ell\right] = \check{\mathbf{P}} \otimes \mathbf{I}_\ell$$

  and evolves according to $\check{\mathbf{T}}^{\mathcal{E}}$ until it enters a phase in $\mathcal{S}_+$.

- When $\widetilde{M}(t)$ hits the upper boundary b, the phase process immediately enters $(\widehat{\mathcal{S}}_0 \times \mathcal{S}_\ell) \cup (\mathcal{S}_- \times \mathcal{S}_\ell)$ according to the matrix

$$\widehat{\mathbf{P}}^{\mathcal{E}} = \left[\widehat{\mathbf{P}}_{-0} \otimes \mathbf{I}_\ell \;\middle\vdots\; \widehat{\mathbf{P}}_{-+} \otimes \mathbf{I}_\ell\right] = \widehat{\mathbf{P}} \otimes \mathbf{I}_\ell$$

  and evolves according to $\widehat{\mathbf{T}}^{\mathcal{E}}$ until it enters a phase in $\mathcal{S}_-$.

- We specify $\mathbf{T}^{\mathcal{E}}$, $\check{\mathbf{T}}^{\mathcal{E}}$ and $\widehat{\mathbf{T}}^{\mathcal{E}}$ according to what behaviour we desire from the horizon. Below, we give an example where the horizon is independent to the rest of the phase process.

- If at any stage $\widetilde{\varphi}(t) = 0$, the horizon is reached and the process terminates.

Note that as before, increasing ℓ will increase the size of the matrices which represent the model, which can slow down computations.

6.2.3 Example: independent energy process

As an example, suppose that the energy process is completely independent of the phase process $\widetilde{\varphi}(t)$. Let us also assume from here on that there are no phases $i \in \mathcal{S}$ with $c_i = 0$, i.e., $\mathcal{S}_0 = \emptyset$. Then we can construct the infinitesimal generator of $\widetilde{\varphi}(t)$ in the same manner as $\mathbf{T}^{\#}$ was constructed in Section 5.1.2. We denote this new generator $\mathbf{T}^{\mathcal{E}}$. Since $\mathcal{S}_0 = \emptyset$,

$$\mathbf{T}^{\mathcal{E}} = \begin{bmatrix} \mathbf{T}^{\mathcal{E}}_{++} & \mathbf{T}^{\mathcal{E}}_{+-} \\ \mathbf{T}^{\mathcal{E}}_{-+} & \mathbf{T}^{\mathcal{E}}_{--} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{++} \oplus \mathbf{K} & \mathbf{T}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}_{--} \oplus \mathbf{K} \end{bmatrix} = \mathbf{T} \oplus \mathbf{K}.$$

We can construct $\widehat{\mathbf{T}}^{\mathcal{E}}$ and $\check{\mathbf{T}}^{\mathcal{E}}$, the respective generators at the upper and lower boundaries in a similar fashion:

$$\widehat{\mathbf{T}}^{\mathcal{E}} = \left[\widehat{\mathbf{T}}^{\mathcal{E}}_{00} \;\middle\vdots\; \widehat{\mathbf{T}}^{\mathcal{E}}_{0-}\right] = \left[\widehat{\mathbf{T}}_{00} \oplus \mathbf{K} \;\middle\vdots\; \widehat{\mathbf{T}}_{0-} \otimes \mathbf{I}_\ell\right],$$
$$\check{\mathbf{T}}^{\mathcal{E}} = \left[\check{\mathbf{T}}^{\mathcal{E}}_{00} \;\middle\vdots\; \check{\mathbf{T}}^{\mathcal{E}}_{0+}\right] = \left[\check{\mathbf{T}}_{00} \oplus \mathbf{K} \;\middle\vdots\; \check{\mathbf{T}}_{0+} \otimes \mathbf{I}_\ell\right].$$

Once again, the matrix entries containing a '$\oplus\,\mathbf{K}$' term denote transitions either in $\mathcal{S}$ *or* $\mathcal{S}_\ell$. The elements with a '$\otimes\,\mathbf{I}_\ell$' term denote transitions in $\mathcal{S}$ only.

We now have a fully functioning bounded stochastic fluid model $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0}$, defined by the matrices $\mathbf{C}^{\mathcal{E}}$, $\widehat{\mathbf{P}}^{\mathcal{E}}$, $\check{\mathbf{P}}^{\mathcal{E}}$, $\mathbf{T}^{\mathcal{E}}$, $\widehat{\mathbf{T}}^{\mathcal{E}}$ and $\check{\mathbf{T}}^{\mathcal{E}}$. Then the process $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0}$ has the same first passage probabilities (and thus the same passage *times*) as the process $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$ terminated at time H.

Time dependent quantities

This model is identical in concept to the model introduced in Section 5.1, in that there is a level and phase process which evolves over time, as well as a horizon which terminates the process when reached. Therefore, we are able to derive time-dependent LSTs for the model ($\widehat{\mathbf{\Psi}}(s)$, $\widehat{\mathbf{\Xi}}(s)$, $\widehat{\mathbf{G}}(s;z)$, $\widehat{\mathbf{H}}(s;z)$, ${}_{x}^{y}\widehat{\mathbf{G}}(s;z)$, ${}_{x}^{y}\widehat{\mathbf{H}}(s;z)$) in the exact same manner.

Recall from Section 5.1.3 (in particular, see Killed hitting times) that these quantities are 'imbued' with the horizon, that is, they account for the fact that if the process is terminated, a path may never reach its destination.

Sadly, this is not quite the fluid model we are looking for: we want a *dependent* energy process (namely, one that is dependent on the dominant symbiont genus).

6.2.4 Moving to a multi-threshold model

We have successfully introduced one of the features we set out to in Section 6.1, but are yet to incorporate the two thresholds ζ_C and ζ_D. To do this, we need to think about how we want the process to evolve on either side of these thresholds. More explicitly, we want to incorporate the following behaviour:

- In the interval $(\zeta_D, b]$, the energy buffer should be increasing in all phases.
- In the interval (ζ_C, ζ_D), the energy buffer should be increasing when C is the dominant genus, and decreasing when D is the dominant genus.
- In the interval $[0, \zeta_C)$, the energy buffer should be decreasing in all phases.
- The energy buffer should persist across visits to each threshold.

The last feature is perhaps the most important: in some sense it will allow the model to keep a record of past bleaching events; however, incorporating this will require some extra machinery, which we introduce in the next section.

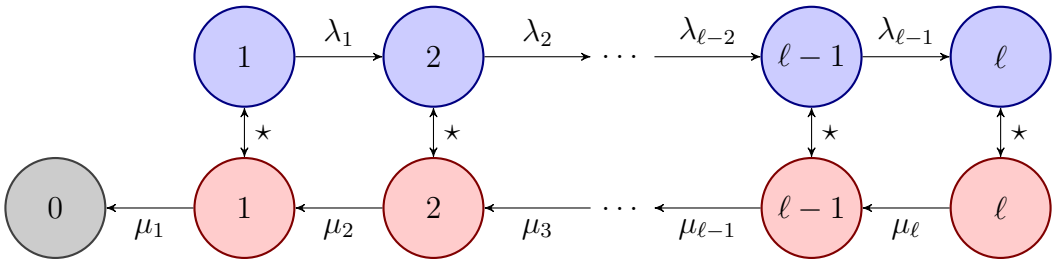


Figure 6.2: A visualisation of the horizon H as a CTMC, where the behaviour is modulated ($\star$) between increasing ($\widehat{\mathbf{K}}$) and decreasing ($\check{\mathbf{K}}$).

6.3 A level-dependent model

In this section, we attempt to incorporate two thresholds into the energy model. To do this, we look towards the matrix $\mathbf{K}$, which defines the horizon process. Note that $\mathbf{K}$ can be thought of as the sum of two matrices: $\mathbf{K} = \widehat{\mathbf{K}} + \check{\mathbf{K}}$, where

$$\widehat{\mathbf{K}} = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & -\lambda_2 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_{\ell-1} & \lambda_{\ell-1} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \tag{6.1}$$

$$\check{\mathbf{K}} = \begin{bmatrix} -\mu_1 & 0 & 0 & \dots & 0 & 0 \\ \mu_2 & -\mu_2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\mu_{\ell-1} & 0 \\ 0 & 0 & 0 & \dots & \mu_\ell & -\mu_\ell \end{bmatrix}. \tag{6.2}$$

The matrix $\widehat{\mathbf{K}}$ represents the *birth* part of the horizon, containing transitions from each state i to $i+1$. The matrix $\check{\mathbf{K}}$ represents the *death* part of the horizon, containing transitions from each state j to $j-1$.

Note that $\widehat{\mathbf{K}}$ is a conservative generator, whereas $\check{\mathbf{K}}$ (and hence $\mathbf{K}$) is not.

Depending on the behaviour desired, we may set each λ_i and μ_i to be any positive number, but henceforth, we set $\lambda_i = \lambda$ and $\mu_i = \mu$ for simplicity.

Suppose that instead of the horizon evolving as a regular PH process as in Figure 6.1, it was modulated between a process that evolved according to $\widehat{\mathbf{K}}$ and a process that evolved according to $\check{\mathbf{K}}$, as illustrated in Figure 6.2.

If the modulation transitions ($\star$) between increasing and decreasing states were exponentially distributed, then this would be a standard PH process. Instead, we look to modulate the transitions by the fluid process from Section 6.2.

6.3.1 Partitioning by dominant genus

Our goal is to partition the process $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0}$ in two ways: by making a distinction between when C or D is the dominant genus, and by the position of the level in relation to the thresholds ζ_C and ζ_D. The former is done entirely within the phase process as follows, and builds on the existing partition of $\mathcal{S}$ into $\mathcal{S}_+$ and $\mathcal{S}_-$.

We further partition $\mathcal{S}$ according to whether C or D is the dominant genus, and apply the same partition to matrices $\mathbf{T}$ and $\mathbf{C}$:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}^{c}_{++} & \mathbf{T}^{cd}_{++} & \mathbf{T}^{c}_{+-} & \mathbf{T}^{cd}_{+-} \\ \mathbf{T}^{dc}_{++} & \mathbf{T}^{d}_{++} & \mathbf{T}^{dc}_{+-} & \mathbf{T}^{d}_{+-} \\ \mathbf{T}^{c}_{-+} & \mathbf{T}^{cd}_{-+} & \mathbf{T}^{c}_{--} & \mathbf{T}^{cd}_{--} \\ \mathbf{T}^{dc}_{-+} & \mathbf{T}^{d}_{-+} & \mathbf{T}^{dc}_{--} & \mathbf{T}^{d}_{--} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}^{c}_{+} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{d}_{+} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{c}_{-} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}^{d}_{-} \end{bmatrix},$$

where a 'c' or 'd' superscript indicates that C or D is the dominant genus, and 'cd' indicates a change in dominant genus from C to D (vice versa for 'dc').

In the basic model, we only included two phases (growth and bleaching) for each genus, and so every sub-matrix here is scalar; however, to allow for more complicated behaviour to be incorporated, we retain the matrix notation.

6.3.2 Partitioning by level

Next, we partition the process $\{\widetilde{\mathcal{E}}(t)\}_{t\geq 0}$ according to the position of the level in relation to the thresholds ζ_C and ζ_D. This is more involved than the previous partition, since we must now take the level into account, and not just the phase. Herein lies the 'extra machinery' mentioned at the end of Section 6.2.

Partition the interval $[0, b]$ into three sub-intervals, $[0, \zeta_C)$, (ζ_C, ζ_D) and $(\zeta_D, b]$. We then define a fluid process for each of these sub-intervals; the full process is called a *level-dependent* fluid model. The behaviour at ζ_C and ζ_D will be explained in more detail briefly, but boils down to switching from one model into another.

To define the full fluid model, we specify matrices $\mathbf{C}$, $\widehat{\mathbf{P}}$, $\check{\mathbf{P}}$, $\mathbf{T}$, $\widehat{\mathbf{T}}$ and $\check{\mathbf{T}}$ for each interval. Fortunately, not all matrices are required for all models: in the upper interval, we do not require $\check{\mathbf{P}}$ or $\check{\mathbf{T}}$; in the lower interval we do not require $\widehat{\mathbf{P}}$ or $\widehat{\mathbf{T}}$; and in (ζ_C, ζ_D), we require $\mathbf{C}$ and $\mathbf{T}$ only.

We denote these models as Model 1, Model 2 and Model 3, for the intervals $[0, \zeta_C)$, (ζ_C, ζ_D) and $(\zeta_D, b]$, respectively, and define the fluid matrices as follows.

Level-dependent matrices

Let $\mathbf{C}^{(1)} = \mathbf{C}^{(2)} = \mathbf{C}^{(3)} = \mathbf{C}^{\mathcal{E}}$. Then, define the infinitesimal generators:

$$\mathbf{T}^{(1)} = \begin{bmatrix} \mathbf{T}^{c}_{++} \oplus \check{\mathbf{K}} & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{++} \oplus \check{\mathbf{K}} & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{c}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{--} \oplus \check{\mathbf{K}} & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{--} \oplus \check{\mathbf{K}} \end{bmatrix},$$

$$\mathbf{T}^{(2)} = \begin{bmatrix} \mathbf{T}^{c}_{++} \oplus \widehat{\mathbf{K}} & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{++} \oplus \check{\mathbf{K}} & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{c}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{--} \oplus \widehat{\mathbf{K}} & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{--} \oplus \check{\mathbf{K}} \end{bmatrix},$$

$$\mathbf{T}^{(3)} = \begin{bmatrix} \mathbf{T}^{c}_{++} \oplus \widehat{\mathbf{K}} & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{++} \oplus \widehat{\mathbf{K}} & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{c}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{c}_{--} \oplus \widehat{\mathbf{K}} & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^{d}_{--} \oplus \widehat{\mathbf{K}} \end{bmatrix}.$$

In addition, define the boundary infinitesimal generators:

$$\check{\mathbf{T}}^{(1)} = \left[\check{\mathbf{T}}_{00} \oplus \check{\mathbf{K}} \;\middle|\; \check{\mathbf{T}}_{0+} \otimes \mathbf{I}_\ell\right],$$
$$\widehat{\mathbf{T}}^{(3)} = \left[\widehat{\mathbf{T}}_{00} \oplus \widehat{\mathbf{K}} \;\middle|\; \widehat{\mathbf{T}}_{0-} \otimes \mathbf{I}_\ell\right].$$

The boundary probability matrices are simply $\check{\mathbf{P}}^{(1)} = \check{\mathbf{P}}^{\mathcal{E}}$ and $\widehat{\mathbf{P}}^{(3)} = \widehat{\mathbf{P}}^{\mathcal{E}}$.

Define $\{\overline{\mathcal{E}}(t)\}_{t\geq 0} = \{(\overline{M}(t), \overline{\varphi}(t)) : t \geq 0\}$ as a bounded, level-dependent fluid model, with level $\overline{M}(t) \in [0, b]$ and phase process $\overline{\varphi}(t) \in \mathcal{S}^{\mathcal{E}}$ as follows:

- When $\overline{M}(t) \in (0, \zeta_C)$, the infinitesimal generator of the phase process is $\mathbf{T}^{(1)}$, and the level evolves according to rates in $\mathbf{C}^{(1)} = \mathbf{C}^{\mathcal{E}}$.
- When $\overline{M}(t) \in (\zeta_C, \zeta_D)$, the infinitesimal generator of the phase process is $\mathbf{T}^{(2)}$, and the level evolves according to rates in $\mathbf{C}^{(2)} = \mathbf{C}^{\mathcal{E}}$.
- When $\overline{M}(t) \in (\zeta_D, b)$, the infinitesimal generator of the phase process is $\mathbf{T}^{(3)}$, and the level evolves according to rates in $\mathbf{C}^{(3)} = \mathbf{C}^{\mathcal{E}}$.

At the boundaries, we specify the following behaviour:

- When $\overline{M}(t)$ hits level 0 in a downwards phase, the process transitions to a phase in $(\breve{\mathcal{S}} \times \mathcal{S}_\ell) \cup (\mathcal{S}_+ \times \mathcal{S}_\ell)$ according to $\breve{\mathbf{P}}^{(1)}$. If the phase process enters $\breve{\mathcal{S}} \times \mathcal{S}_\ell$ the phase process evolves according to the generator $\breve{\mathbf{T}}^{(1)}$ until the process transitions to a phase in $\mathcal{S}_+ \times \mathcal{S}_\ell$.
- When $\overline{M}(t)$ hits level b in an upwards phase, the process transitions to a phase in $(\widehat{\mathcal{S}} \times \mathcal{S}_\ell) \cup (\mathcal{S}_- \times \mathcal{S}_\ell)$ according to $\widehat{\mathbf{P}}^{(3)}$. If the phase process enters $\widehat{\mathcal{S}} \times \mathcal{S}_\ell$ the phase process evolves according to the generator $\widehat{\mathbf{T}}^{(3)}$ until the process transitions to a phase in $\mathcal{S}_- \times \mathcal{S}_\ell$.

In regard to the behaviour at ζ_C and ζ_D, observe that $\mathbf{T}^{(1)}$, $\mathbf{T}^{(2)}$ and $\mathbf{T}^{(3)}$ have the same dimension, and each row and column correspond to the same state in $\mathcal{S}^{\mathcal{E}}$. Thus, at the thresholds ζ_C and ζ_D, the process swaps to a different set of matrices, and remains in the same phase. For example, consider a path which begins at level b with generator $\mathbf{T}^{(3)}$ and hits level ζ_D in phase (i_1, i_2). Now $\overline{M}(t) \in (\zeta_C, \zeta_D)$, so $\overline{\varphi}(t)$ evolves with generator $\mathbf{T}^{(2)}$, and is still in phase (i_1, i_2).

Interpreting $\{\overline{\mathcal{E}}(t)\}_{t \geq 0}$

In essence, the process $\{\overline{\mathcal{E}}(t)\}_{t \geq 0}$ is the same as the process $\{\widetilde{\mathcal{E}}(t)\}_{t \geq 0}$, but the behaviour of the horizon changes at ζ_C and ζ_D. Understanding how and why this behaviour changes is critical to understanding the energy model.

At a glance, the matrices $\mathbf{T}^{(1)}$, $\mathbf{T}^{(2)}$ and $\mathbf{T}^{(3)}$ are exactly the same; the difference is that they contain different combinations of the matrices $\widehat{\mathbf{K}}$ and $\breve{\mathbf{K}}$ on the diagonals. Notice that $\mathbf{T}^{(1)}$ contains $\breve{\mathbf{K}}$ only, $\mathbf{T}^{(3)}$ contains $\widehat{\mathbf{K}}$ only, and $\mathbf{T}^{(2)}$ contains both $\widehat{\mathbf{K}}$ and $\breve{\mathbf{K}}$. Importantly, the diagonal entries of $\mathbf{T}^{(2)}$ corresponding to genus C contain $\widehat{\mathbf{K}}$, whereas the diagonal entries corresponding to genus D contain $\breve{\mathbf{K}}$.

Recall from the introduction to Section 6.3 that $\widehat{\mathbf{K}}$ represents the *birth* part of the horizon, whereas the matrix $\breve{\mathbf{K}}$ represents the *death* part of the horizon. By taking the Kronecker sum with $\widehat{\mathbf{K}}$, we allow the horizon to increase only; by taking the Kronecker sum with $\breve{\mathbf{K}}$, we allow the horizon to decrease only. Thus:

- In the interval $[0, \zeta_C)$, the energy stage decreases only;
- In the interval (ζ_C, ζ_D), the energy stage can only increase when C is the dominant genus, and can only decrease when D is the dominant genus;
- In the interval $(\zeta_D, b]$, the energy stage increases only; and
- The energy stage persists across visits to each threshold.

Figure 6.3 illustrates how the energy stage evolves in each interval.

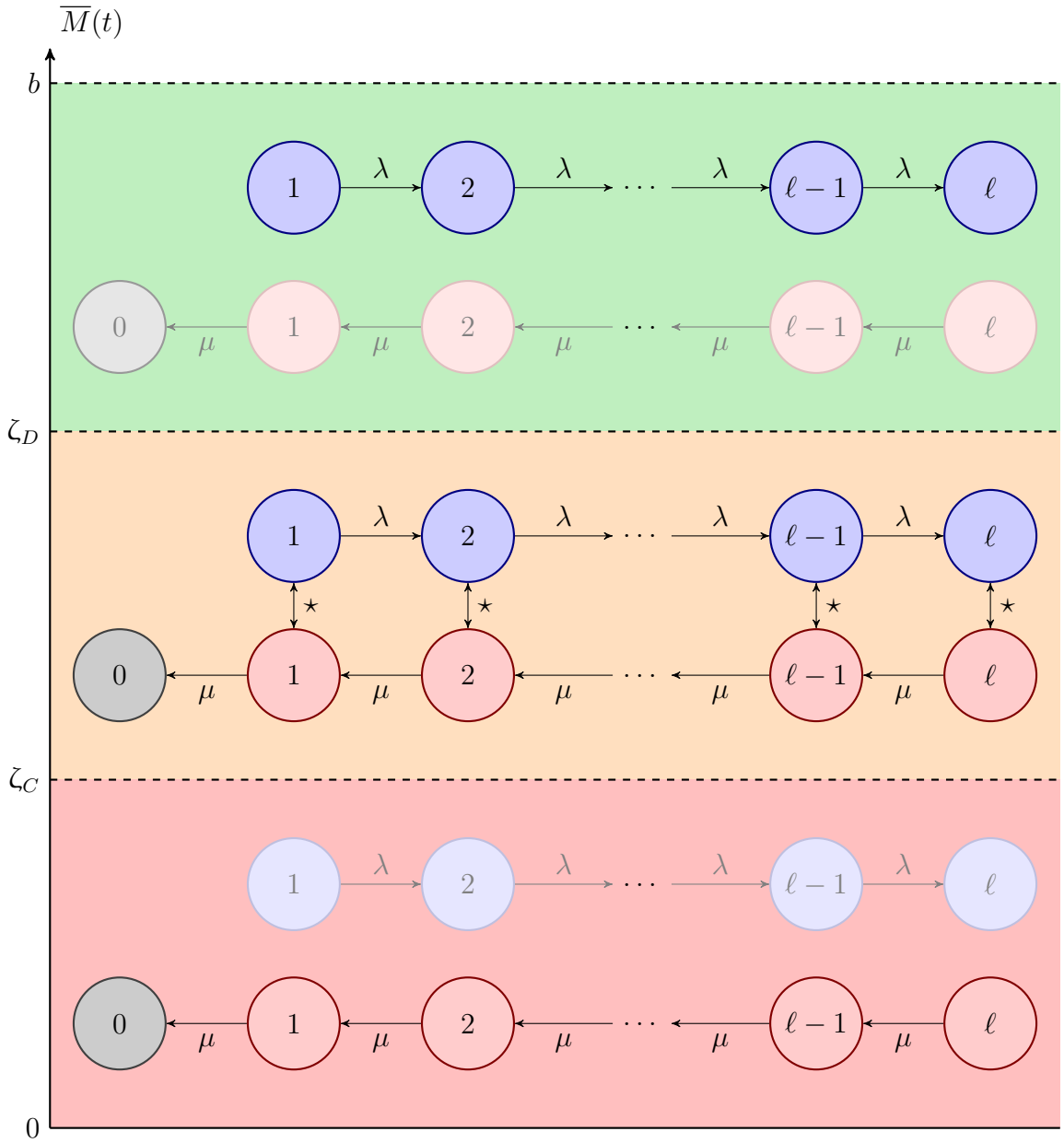


Figure 6.3: A visualisation of the *energy stage* component of the phase process $\overline{\varphi}(t)$ in the level-dependent fluid model $\{\overline{\mathcal{E}}(t)\}_{t\geq 0} = \{(\overline{M}(t), \overline{\varphi}(t)) : t \geq 0\}$. The behaviour of the energy stage differs based on the location of the level $\overline{M}(t)$ in relation to the intervals $[0, \zeta_C)$, (ζ_C, ζ_D) and $(\zeta_D, b]$. Transitions between intervals occur when the level $\overline{M}(t)$ hits either ζ_C or ζ_D. Faded states indicate that the phase never transitions to that state while in the specified interval. In the interval (ζ_C, ζ_D), the process is modulated ($\star$) between increasing and decreasing according to whether genus C or genus D is dominant.

Different energy rates

Currently, genus C and genus D have the same energy production rate, since they both utilise the same $\widehat{\mathbf{K}}$ and $\breve{\mathbf{K}}$ matrices. To allow for different energy production rates, we may define matrices $\widehat{\mathbf{K}}^c$, $\breve{\mathbf{K}}^c$, $\widehat{\mathbf{K}}^d$ and $\breve{\mathbf{K}}^d$, where:

$$\widehat{\mathbf{K}}^x = \begin{bmatrix} -\lambda_x & \lambda_x & 0 & \dots & 0 & 0 \\ 0 & -\lambda_x & \lambda_x & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_x & \lambda_x \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

$$\breve{\mathbf{K}}^x = \begin{bmatrix} -\mu_x & 0 & 0 & \dots & 0 & 0 \\ \mu_x & -\mu_x & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\mu_x & 0 \\ 0 & 0 & 0 & \dots & \mu_x & -\mu_x \end{bmatrix},$$

for $x \in \{c, d\}$. To account for the fact that genus C produces energy faster than genus D, we require that $\lambda_c > \lambda_d$ and $\mu_c < \mu_d$. Define the new generators:

$$\mathbf{T}^{(1*)} = \begin{bmatrix} \mathbf{T}^c_{++} \oplus \breve{\mathbf{K}}^c & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{++} \oplus \breve{\mathbf{K}}^d & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^c_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{--} \oplus \breve{\mathbf{K}}^c & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{--} \oplus \breve{\mathbf{K}}^d \end{bmatrix},$$

$$\mathbf{T}^{(2*)} = \begin{bmatrix} \mathbf{T}^c_{++} \oplus \widehat{\mathbf{K}}^c & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{++} \oplus \breve{\mathbf{K}}^d & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^c_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{--} \oplus \widehat{\mathbf{K}}^c & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{--} \oplus \breve{\mathbf{K}}^d \end{bmatrix},$$

$$\mathbf{T}^{(3*)} = \begin{bmatrix} \mathbf{T}^c_{++} \oplus \widehat{\mathbf{K}}^c & \mathbf{T}^{cd}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{++} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{++} \oplus \widehat{\mathbf{K}}^d & \mathbf{T}^{dc}_{+-} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{+-} \otimes \mathbf{I}_\ell \\ \mathbf{T}^c_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{cd}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^c_{--} \oplus \widehat{\mathbf{K}}^c & \mathbf{T}^{cd}_{--} \otimes \mathbf{I}_\ell \\ \mathbf{T}^{dc}_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{-+} \otimes \mathbf{I}_\ell & \mathbf{T}^{dc}_{--} \otimes \mathbf{I}_\ell & \mathbf{T}^d_{--} \oplus \widehat{\mathbf{K}}^d \end{bmatrix}.$$

In addition, define the new boundary infinitesimal generators:

$$\breve{\mathbf{T}}^{(1*)} = \left[\breve{\mathbf{T}}_{00} \oplus \breve{\mathbf{K}} \;\middle|\; \breve{\mathbf{T}}_{0+} \otimes \mathbf{I}_\ell\right] = \breve{\mathbf{T}}^{(1)},$$
$$\widehat{\mathbf{T}}^{(3*)} = \left[\begin{array}{cc|cc} \widehat{\mathbf{T}}^c_{00} \oplus \widehat{\mathbf{K}}^c & \widehat{\mathbf{T}}^{dc}_{00} \otimes \mathbf{I}_\ell & \widehat{\mathbf{T}}^c_{0-} \otimes \mathbf{I}_\ell & \widehat{\mathbf{T}}^{cd}_{0-} \otimes \mathbf{I}_\ell \\ \widehat{\mathbf{T}}^{dc}_{00} \otimes \mathbf{I}_\ell & \widehat{\mathbf{T}}^d_{00} \oplus \widehat{\mathbf{K}}^d & \widehat{\mathbf{T}}^{dc}_{0-} \otimes \mathbf{I}_\ell & \widehat{\mathbf{T}}^d_{0-} \otimes \mathbf{I}_\ell \end{array}\right].$$

We do not change the generator at the lower boundary, since when the algal density is zero, neither genus is dominant. Thus, we must also specify a separate $\breve{\mathbf{K}}$ matrix for the lower boundary. For simplicity, we specify $\breve{\mathbf{K}}$ as in Eq. (6.2) with the single parameter $\mu = \max\{\mu_c, \mu_d\} = \mu_d$.

Thus, if we set $\mathbf{T}^{(1)} = \mathbf{T}^{(1*)}$, $\mathbf{T}^{(2)} = \mathbf{T}^{(2*)}$, $\mathbf{T}^{(3)} = \mathbf{T}^{(3*)}$, and $\widehat{\mathbf{T}}^{(3)} = \widehat{\mathbf{T}}^{(3*)}$, our level-dependent process evolves in exactly the same way, but the energy buffer increases at a higher rate and decreases at a lower rate when genus C is dominant.

Flexibility

The energy model framework allows for much more flexibility than we are taking advantage of: we are able to define different $\mathbf{T}$ and $\mathbf{C}$ matrices for each interval, such that the phase transition rates and fluid input rates are also level dependent. For example, we could use different $\mathbf{C}$ matrices to (crudely) approximate logistic algal growth rates. We could also specify a more complicated horizon process, rather than simply using λ_x and μ_x. To truly push the limits of the energy model, the framework could even be generalised to $n \in \mathbb{N}$ thresholds, although this would make for a much more complicated mortality process.

Note on level-dependent time-dependent quantities

In what follows, we make use of several time-dependent quantities derived from the model $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$. We only consider paths between the intervals $[0, \zeta_C)$, (ζ_C, ζ_D) and $(\zeta_D, b]$, and thus each quantity belongs to one of Model 1, Model 2 or Model 3. This will be denoted both by colour, as well as scripts to the left of the quantity, to make it clear what model the quantity belongs to—for instance, x and y in ${}^{y}_{x}\widehat{\mathbf{H}}_{-+}(s;z)$ indicate that the quantity evolves in the interval (x, y).

To calculate each quantity, we simply use the fluid matrices which define the corresponding model in our usual fluid model construction (that is, the construction in Section 2.4). For example, in the case of ${}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{H}}_{-+}(s;\zeta_D)$, we construct the matrices $\widehat{\mathbf{\Psi}}(s)$, $\widehat{\mathbf{\Xi}}(s)$ by solving Riccati equations involving $\mathbf{T}^{(2)}$ and $\mathbf{C}^{(2)}$; we then use these to construct $\widehat{\mathbf{G}}(s;z)$, $\widehat{\mathbf{H}}(s;z)$ and so on.

6.3.3 Time to mortality in the energy model

We now attempt to derive an expression for the LST of the TTM in the energy model, using an approach very similar to Section 5.2.2, however, we will encounter additional difficulties this time around. We begin with the following lemma, which is analogous to Lemma 5.3, and use this to construct the TTM as in Theorem 5.5

Lemma 6.1. *Suppose the fluid process* $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$ *hits level* ζ_D *from above in phase* $\boldsymbol{i} \in \mathcal{S}_- \times \mathcal{S}_\ell$*. Then the LST of the time to return to* ζ_D *in phase* $\boldsymbol{j} \in \mathcal{S}_+ \times \mathcal{S}_\ell$*, before the horizon* H *is reached, is* $[\overline{\mathbf{D}}(s)]_{\boldsymbol{i},\boldsymbol{j}}$*, where*

$$\begin{aligned}\overline{\mathbf{D}}(s) = {}& {}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{H}}_{-+}(s;\zeta_D) + {}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{G}}_{--}(s;\zeta_D)\,{}^{\zeta_C}_{0}\widetilde{\boldsymbol{\Xi}}(s) \\ & \times \left(\mathbf{I} - {}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{G}}_{+-}(s;\zeta_C)\,{}^{\zeta_C}_{0}\widetilde{\boldsymbol{\Xi}}(s)\right)^{-1} {}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{H}}_{++}(s;\zeta_C).\end{aligned} \tag{6.3}$$

Proof. Consider a path starting at ζ_D in phase $\boldsymbol{i} \in \mathcal{S}_- \times \mathcal{S}_\ell$ and ending at ζ_D in phase $\boldsymbol{j} \in \mathcal{S}_+ \times \mathcal{S}_\ell$ at some time prior to H. This can occur in two ways:

1. The level process returns to ζ_D in taboo of ζ_C and prior to H. The LST of the time taken on this journey is ${}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{H}}_{-+}(s;\zeta_D)$.

2. The level hits ζ_C before returning to ζ_D. The LST of the time taken on this journey is ${}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{G}}_{--}(s;\zeta_D)$.

   - After the level process hits ζ_C from above, the level must return to ζ_C. The LST of the time taken on this journey is ${}^{\zeta_C}_{0}\widetilde{\boldsymbol{\Xi}}(s)$.
   - After the level process returns to ζ_C from below, it can make any number of journeys back to ζ_C in the interval (ζ_C, ζ_D), followed by a return journey to ζ_C in the interval $[0, \zeta_C)$. By the convolution theorem, the LST of the time taken to complete one such journey is ${}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{G}}_{+-}(s;\zeta_C)\,{}^{\zeta_C}_{0}\widetilde{\boldsymbol{\Xi}}(s)$, since these journeys are conditionally independent, and thus the LST of the time taken to complete *any* number of such journeys is $\left(\mathbf{I} - {}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{G}}_{+-}(s;\zeta_C)\,{}^{\zeta_C}_{0}\widetilde{\boldsymbol{\Xi}}(s)\right)^{-1}$.
   - The level process must make one journey from ζ_C to ζ_D, in taboo of ζ_C (otherwise such a path would be included in the previous LST). The LST of the time taken on this journey is ${}^{\zeta_D}_{\zeta_C}\widehat{\mathbf{H}}_{++}(s;\zeta_C)$.

   All of this must happen before the horizon H is reached, which is why we record the energy stage at all points of the journey.

Hence, the LST of the total time elapsed on a return journey to ζ_D, starting in phase $\boldsymbol{i}$ and ending in phase $\boldsymbol{j}$, prior to H, is the $(\boldsymbol{i}, \boldsymbol{j})$ element of $\overline{\mathbf{D}}(s)$.

■

Theorem 6.2. *The LST of the TTM under the level-dependent model* $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$*, given that the process starts in phase* $\boldsymbol{i} \in \mathcal{S}_- \times \mathcal{S}_\ell$ *at level* b*, is*

$$\widehat{F}^{\mathcal{E}}_{\mathrm{TTM}}(s) = \boldsymbol{e}_{\boldsymbol{i}}^{\intercal}\, {}_{\zeta_D}^{b}\widetilde{\mathbf{G}}_{--}(s;b)\left(\mathbf{I} - \overline{\mathbf{D}}(s)\, {}_{\zeta_D}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right)^{-1} \boldsymbol{\Omega}^{\mathcal{E}}(s), \tag{6.4}$$

where $\overline{\mathbf{D}}(s)$ *is defined as in Lemma 6.1, and* $\boldsymbol{\Omega}^{\mathcal{E}}(s)$ *is the LST of the time until the horizon is reached on a journey below* ζ_D*, given that the horizon is reached (mortality occurs) before returning to* ζ_D*.*

Proof. Consider a path which begins at level b in phase $\boldsymbol{i} \in \mathcal{S}_- \times \mathcal{S}_\ell$ and eventually reaches the cemetery state 0, at which point mortality occurs:

- For mortality to occur, the level must first drain from b to ζ_D, possibly returning to b before reaching ζ_D. The LST of the time taken for this journey is ${}_{\zeta}^{b}\widetilde{\mathbf{G}}_{--}(s;b)$. Since the initial phase is $\boldsymbol{i} \in \mathcal{S}_-$, we pre-multiply by the (row) unit vector $\boldsymbol{e}_{\boldsymbol{i}}^{\intercal}$, which only takes the $\boldsymbol{i}^{\text{th}}$ row of this matrix.

- Once the level hits ζ_D, the process can make any number of journeys below ζ_D, followed by a journey above ζ_D, as long as the journey below ζ_D returns to ζ_D before the horizon is reached.

  The LST of the journey below ζ_D is $\overline{\mathbf{D}}(s)$ from Lemma 6.1, and the LST of the journey above ζ_D is ${}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)$. Thus, the LST of the full journey is $\overline{\mathbf{D}}(s)\, {}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)$ by the convolution theorem (since the times for each segment are conditionally independent). Therefore, the LST of any number of such journeys is $\sum_{m=0}^{\infty}[\overline{\mathbf{D}}(s)\, {}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)]^m = \left(\mathbf{I} - \overline{\mathbf{D}}(s)\, {}_{\zeta}^{b}\widetilde{\boldsymbol{\Psi}}(s)\right)^{-1}$.

- The process must eventually make a single journey below ζ_D which ends in mortality. The LST of the time spent on this journey before the instant of mortality is $\boldsymbol{\Omega}^{\mathcal{E}}(s)$. *Note:* $\boldsymbol{\Omega}^{\mathcal{E}}(s)$ *has not yet been specified.*

Once again, since these paths are conditionally independent, the convolution theorem applies, and hence the result follows.

■

But what about $\boldsymbol{\Omega}^{\mathcal{E}}(s)$?

To complete the expression for the TTM, we require the final component, $\boldsymbol{\Omega}^{\mathcal{E}}(s)$. Unfortunately, it is not quite as easy to obtain $\boldsymbol{\Omega}^{\mathcal{E}}(s)$ as it was to obtain $\boldsymbol{\Omega}(s)$ in Section 5.2.2, which relied on the fact that the horizon process and $\{\widetilde{\mathcal{F}}(t)\}_{t\geq 0}$ were independent. In this case, the horizon process and fluid process $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$ are *not independent*, since the behaviour of the horizon depends on the level and phase.

Since we cannot derive $\mathbf{\Omega}^{\mathcal{E}}(s)$ in the same manner as $\mathbf{\Omega}(s)$, one wonders if we can derive $\mathbf{\Omega}^{\mathcal{E}}(s)$ *at all*. Unfortunately, we do not provide an answer to that question here, other than to say that we have not found an efficient expression for $\mathbf{\Omega}^{\mathcal{E}}(s)$, and even if we *could* find such an expression, computing and inverting the LST of the TTM would be incredibly computationally expensive for a large number of horizon stages. The energy model is even more complex than the model in Section 5.2, and as a result, computations will be at least as slow.

For this reason, we once again look to the alternative: simulation. Although obtaining an analytic expression for the TTM has proven difficult, we can still simulate the energy model, which is where we head next.

6.4 Single-genus (SG) models

The energy model is the most realistic model that we cover in this thesis, and so it is the best way to address our research questions from Section 2.1.5. In the remainder of this chapter, we focus on two questions:

Q1. Under the energy model, does the coral host benefit from hosting both genus C and genus D, or is it more beneficial to host a single genus only?

Q2. How sensitive is the energy model to its parameters? That is, how does the answer to Q1 change if the model parameters are reasonably perturbed?

To answer these questions, we first need to introduce some new machinery. Namely, we need to specify what it means for the coral to host a single genus only, as opposed to both genus C and genus D (as assumed in the energy model).

To achieve this, we specify Models CD, C and D as follows:

- Model CD: The level-dependent model $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$.
- Model C: The level-dependent model $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$ where C is the only genus.
- Model D: The level-dependent model $\{\overline{\mathcal{E}}(t)\}_{t\geq 0}$ where D is the only genus.

To specify the single-genus (SG) models Model C and Model D, we need to specify matrices $\mathbf{C}$, $\widehat{\mathbf{P}}$, $\check{\mathbf{P}}$, $\mathbf{T}$, $\widehat{\mathbf{T}}$ and $\check{\mathbf{T}}$ for each interval. It is easy to define $\mathbf{C}$, $\widehat{\mathbf{P}}$ and $\check{\mathbf{P}}$, since these are simply inherited from Model CD, as follows:

- In Model C:
  - For $i \in \{1, 2, 3\}$, the matrix $\mathbf{C}^{(i)}$ is the sub-matrix of $\mathbf{C}^{(i)}$ under Model CD, only including phases where genus C is dominant;

  - The matrices $\breve{\mathbf{P}}^{(1)}$ and $\widehat{\mathbf{P}}^{(3)}$ are, respectively, subsets of $\breve{\mathbf{P}}^{(1)}$ and $\widehat{\mathbf{P}}^{(3)}$ under Model CD, only including phases where genus C is dominant.

- Similarly for Model D.

In addition, $\mathbf{T}^{(3)} = \mathbf{T}^{(2)}$ under Model C, since both matrices describe the behaviour of the model above ζ_C. Similarly, $\mathbf{T}^{(2)} = \mathbf{T}^{(1)}$ under Model D.

Specifying $\mathbf{T}^{(1)}$, $\breve{\mathbf{T}}^{(1)}$, $\widehat{\mathbf{T}}^{(3)}$ and $\mathbf{T}^{(3)}$ requires some care, since we must determine how the transition rates change when we omit one genus.

Each of the above generators is derived entirely from the generator $\mathbf{T}$. Recall that under the canonical parameters, $\mathbf{T}$ has the form

$$\mathbf{T} = \begin{bmatrix} -(\beta + 2\alpha) & \beta & 2\alpha & 0 \\ 0 & -\alpha & 0 & \alpha \\ \gamma & \delta & -(\gamma + \delta) & 0 \\ \gamma & \delta & 0 & -(\gamma + \delta) \end{bmatrix},$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}^+$. Note that:

- α is the bleaching event rate for genus D, and 2α is the rate for genus C. These rates should stay the same for each SG model.

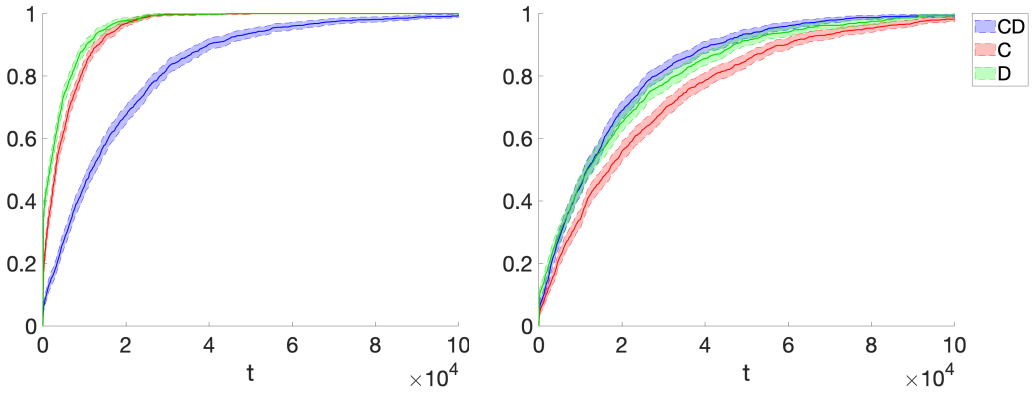
- β is the rate at which genus C is succeeded by genus D. For a SG model, this parameter is unnecessary, and so it should not appear. It is worth noting that removing β alone does not partition the process into genus C and genus D, since the dominant genus can change after recovering from a bleaching event via the parameters γ and δ.

- $\gamma+\delta$ is the rate at which bleaching events end. The probability of genus C or genus D becoming dominant depends on the ratio between γ and δ.

We specify matrices $\mathbf{T}_c$ and $\mathbf{T}_d$ for each SG model, which replace the infinitesimal generator $\mathbf{T}$. Complying with the above statement that the bleaching event rates should be consistent across all models, $\mathbf{T}_c$ and $\mathbf{T}_d$ have the following form:

$$\mathbf{T}_c = \begin{bmatrix} -2\alpha & 2\alpha \\ x & -x \end{bmatrix}, \qquad \mathbf{T}_d = \begin{bmatrix} -\alpha & \alpha \\ y & -y \end{bmatrix},$$

where x and y are, respectively, the recovery rates for genus C and genus D. Choosing values for x and y is somewhat challenging: we consider two methodologies for doing so, which we refer to as SG1 and SG2.



(a) SG1 (b) SG2

Figure 6.4: Simulated CDF of the TTM, for Models CD, C and D, using the starting parameters from Section 6.5 and 1,000 simulations. Shaded areas indicate 95% confidence intervals.

6.4.1 SG1: symbiont stops bleaching

Let the matrices

$$\mathbf{T}_c = \begin{bmatrix} -2\alpha & 2\alpha \\ \gamma & -\gamma \end{bmatrix}, \qquad \mathbf{T}_d = \begin{bmatrix} -\alpha & \alpha \\ \delta & -\delta \end{bmatrix}$$

be the infinitesimal generators of the phase processes for Model C and Model D, respectively. If only genus C is present (Model C), the coral can only recover at rate γ; if only genus D is present (Model D), the coral can only recover at rate δ; when both are present (Model CD), the coral recovers at rate $\gamma + \delta$ as usual.

A possible interpretation of this choice of $\mathbf{T}_c$ and $\mathbf{T}_d$ is that the symbiont genus is actively trying to stop bleaching; when hosting both genera, the host has a higher recovery rate, because both genera are trying to stop bleaching.

Figure 6.4a shows the simulated CDF of the TTM under SG1, for models CD, C and D, for 1,000 simulations, using the starting parameters from Section 6.5. Shaded areas denote the 95% confidence intervals for each value of t plotted. These confidence intervals quickly shrink as the number of simulations increases. Under SG1 in this specification, the CDF of the TTM for Model CD lies far below the CDFs for Models C and D, indicating that the survivability of the coral under Model CD is much higher than its counterparts.

It should be noted that Figure 6.4a only shows the CDF for a single set of parameters, and that under a different set of parameters, the CDF for Model CD could look entirely different. That being said, it is likely that under different sets

of parameters, we will see similar behaviour, since the rate of leaving bleaching phases is much higher in Model CD than in Models C and D.

More importantly, recall from Section 2.1.2 that bleaching is most likely an immune-type response from the coral host. This seems to imply that the coral host has more control over when bleaching ends. This is somewhat inconsistent with SG1, so it may be more appropriate to consider an approach in which the host has more control over bleaching.

6.4.2 SG2: host stops bleaching

Let the matrices

$$\mathbf{T}_c = \begin{bmatrix} -2\alpha & 2\alpha \\ \gamma+\delta & -(\gamma+\delta) \end{bmatrix}, \qquad \mathbf{T}_d = \begin{bmatrix} -\alpha & \alpha \\ \gamma+\delta & -(\gamma+\delta) \end{bmatrix}$$

be the infinitesimal generators of the phase processes for Model C and Model D respectively. Here, the coral recovers at some fixed rate $\gamma+\delta$ in all models. In Model CD, after recovering, the coral is recolonised by genus C with probability $\gamma/(\gamma+\delta)$, and genus D with probability $\delta/(\gamma+\delta)$.

The interpretation of this choice of $\mathbf{T}_c$ and $\mathbf{T}_d$ is that the host is the party attempting to stop bleaching, irrespective of the genus of symbiont currently hosted. This is more consistent with the idea that bleaching is an immune-type response, and that the host has more control over when bleaching ends.

The CDF of the TTM under SG2 is shown in Figure 6.4b, where shaded areas denote the 95% confidence intervals for each value of t. We see that in comparison to Figure 6.4a, the CDFs are much closer; interestingly, Model CD now appears to be the model with the highest probability of mortality at most times.

We use SG2 in the rest of our analysis, since it is more consistent with our understanding of coral bleaching.

6.5 Analysis and results

6.5.1 Parameters

In this section, we analyse the energy model by comparing the single genus models formed under SG2, in an attempt to answer Q1 and Q2. We begin by using the set of canonical parameters from Section 3.2.4, and then vary these parameters to test the sensitivity of the model. First, we need to choose values for the new parameters introduced in the energy model:

- ℓ: the number of energy stages;
- λ_c, λ_d: the rates of energy increase for genus C and D, respectively;
- μ_c, μ_d: the rates of energy decrease for genus C and D, respectively (recall that when the algal density is zero, the rate is $\mu = \max\{\mu_c, \mu_d\} = \mu_d$); and
- ζ_C and ζ_D: the critical thresholds for genus C and D, respectively.

The number of energy stages determines how closely we approximate a continuous energy buffer: more stages yields a better approximation. We start with $\ell = 10$ stages, since this is a large enough value to reasonably approximate a continuous energy buffer. Note: very small values such as $\ell = 2$ lead to very different results, perhaps due to the rigidity of the energy process.

For the energy increase/decrease rates, we use rates $\lambda_d = \mu_d = \ell/\tau$. This way, we are consistent with the Erlangization model from Section 5.3, in which $\mu = \ell/\tau$. Recall that we use the value $\tau = 14$ in the canonical parameters. To account for the fact that genus C produces energy at a rate higher than genus D, we let $\lambda_c = \ell(1+\sigma)/\tau$ and $\mu_c = \ell(1-\sigma)/\tau$, where $\sigma \in [0, 1)$ is a scaling factor representing how much better genus C is at producing energy. For now, we use the value $\sigma = 0.25$.

For the critical thresholds, we use starting values of $\zeta_C = 0.33$ and $\zeta_D = 0.66$.

6.5.2 Methods

To simulate individual paths, we use a variation of Algorithm B.4, with the same modifications at (†) and (‡) as in Chapter 5. To simulate the TTM, we simply 'stitch' together several paths, each of which retains information about the level, phase (including the energy stage), and total time elapsed. More specifically, we simulate an initial path from level b to level ζ_D, and then simulate paths between ζ_C and ζ_D. Not including the initial path, there are four types of possible paths:

- Starting at ζ_C in a downwards phase. We simulate the time until the process returns to ζ_C from below using Model 1.
- Starting at ζ_D in a downwards phase. We simulate the time until the level hits ζ_D (destination) or ζ_C (taboo) using Model 2.
- Starting at ζ_C in an upwards phase. We simulate the time until the level hits ζ_D (destination) or ζ_C (taboo) using Model 2.
- Starting at ζ_D in an upwards phase. We simulate the time until the process returns to ζ_D from above using Model 3.

During each path, the horizon is implicitly tracked by the phase process. Conditional statements placed in Algorithm B.4 at (†) and (‡) check whether the cemetery state has been reached, at which point the simulation terminates. Pseudocode for the full simulation algorithm is shown in Algorithm B.6. The fashionably named `supermodel` structure contains model structures for Model 1, Model 2 and Model 3.

CDFs are cumbersome objects to analyse. Here, we instead use tail probabilities. Recall from Section 4.4.4 that a tail probability is a proxy for the probability that a colony survives a pre-determined amount of time. We can approximate the probability of mortality at some fixed time ω by the value of the simulated CDF at ω. This allows us to compare Models CD, C and D with a single number, rather than an entire CDF.

We use the value $\omega = 5 \times 365 = 1825$ to obtain our tail probabilities, representing five years after the initial bleaching event at $t = 0$. We are interested in the values $F_{CD}(\omega)$, $F_C(\omega)$ and $F_D(\omega)$, which denote the CDFs of the TTM under Models CD, C and D, respectively, each evaluated at $t = \omega$. These are the probabilities that the coral has died within five years of the initial bleaching event.

To obtain $F_{CD}(\omega)$, $F_C(\omega)$ and $F_D(\omega)$, we calculate the proportion of simulations which terminate before ω. For brevity, we henceforth refer to all of these values as F^ω, unless referring to the value under a specific model.

We simulate Models CD, C and D N times each, varying certain parameters to observe whether they have a significant impact on F^ω. Since there are a relatively large number of parameters in this model, we limit our tests to one or two parameters at a time. When varying one parameter, we plot the values of F^ω for each model. When varying two parameters, we use a heatmap to indicate which model produces the highest and lowest values of F^ω.

Using tail probabilities to compare models simplifies the comparison process greatly, but can be prone to misinterpretation. For instance, if each model produces a very similar value of F^ω, with very little variance, then saying that a particular model is the best model is somewhat disingenuous.

To compare models more thoroughly, we use a one-way ANOVA test, which tests whether the models produce significantly different values of F^ω. This allows us to determine whether any comparisons are unsound due to very similar values. We fully explain the ANOVA testing process and further discuss the use of tail probabilities as a comparison method in the insert Comparison methods.

For each comparison, we perform $A = 10$ ANOVA tests and $N = 1000$ simulations per test, per model, resulting in 30,000 simulations for each set of parameters.

Computational shortcuts

When simulating the energy model, we encounter a slight problem: there is technically no upper limit for the time of mortality. For any given time t, it is possible for mortality to occur after t. For example, if the level stays at b the entire time; this is incredibly unlikely, but it is a viable path.

This becomes a particular issue when the probability of mortality becomes very small, as a result of things such as:

- The parameter τ is too large, and thus the process can spend more time below ζ_C and ζ_D without mortality occurring;
- The rate of energy decrease is too small, or the rate of energy increase is too large (the former tends to have a more significant effect) with respect to the mean drift (the mean amount of fluid filled/drained per unit time); or
- The thresholds ζ_C and ζ_D are very small (and so the process spends less time draining energy).

Luckily, there is an easy way to circumvent these issues, which is to terminate a simulation if mortality has not occurred by $\omega + \varepsilon$, for some small $\varepsilon > 0$. In addition, this removes any unnecessary simulation time, since we are not interested in what happens after ω if mortality has not occurred by then.

6.5.3 Results

Effect of ℓ and τ

The parameters ℓ and τ have quite different interpretations in a modelling context; however, they are inextricably linked since the energy increase and decrease rates are proportional to ℓ/τ. We are interested in the effect of both parameters on the probability of mortality before ω in each model.

Figure 6.5 shows the influence of ℓ and τ on F^{ω}, respectively, when the other parameter is fixed (we fix $\tau = 14$ and $\ell = 10$), for all three models. We observe in Figure 6.5a that smaller values of ℓ lead to larger values of F^{ω}, but the overall ordering of the models stays the same for the most part. In Figure 6.5b, we see a roughly linear trend with a negative slope as τ varies. This makes sense, since a larger value of τ should correspond to a smaller probability of mortality, since the rate of energy decrease is lessened.

We repeat the same experiment but with larger values of ℓ in Figure 6.6. In Figure 6.6a, we see that when τ is fixed at 14, there is very little change in F^{ω},

Comparison methods

We use tail probabilities to compare CDFs: each model produces a value F^{ω}, and we use these values to determine which model has the highest and lowest probability of mortality at $t = \omega$.

We should be aware of the drawbacks of using tail probabilities. Namely, that using a single number makes our results prone to misinterpretation, particularly if we are dealing with small numbers. Importantly, *we are losing some detail about coral mortality by focusing on a single point in time.*

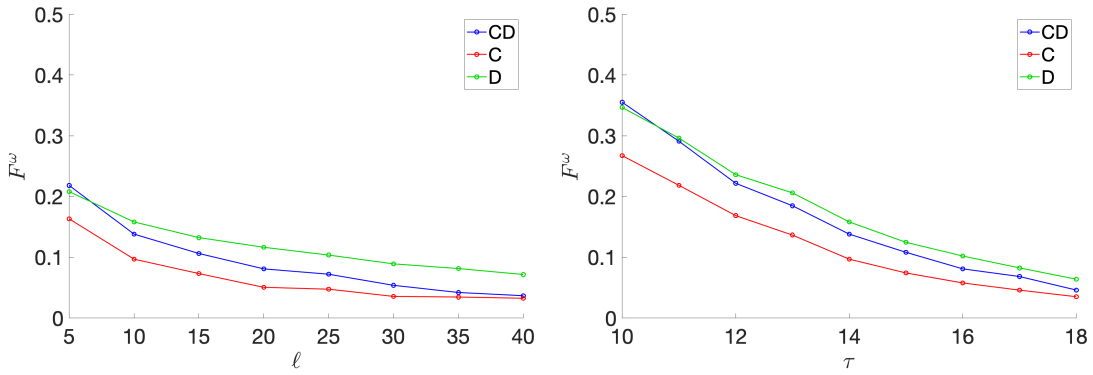
Although we cannot magically glean more information from tail probabilities, we can make an effort to reduce the uncertainty of our results. We do this by performing one-way ANOVA tests on our simulated data, to see whether our results involving F^{ω} are statistically significant. This is done as follows:

- For each set of parameters, simulate A values of F^{ω} under each model. Note that each value of F^{ω} requires N simulations.
- Perform a one-way ANOVA test on the resulting $3 \times A$ data set. The assumption of normality is reasonable, if N is large enough, due to the central limit theorem. The outcome of this test is a p-value indicating the probability that the means of each distribution of F^{ω} are the same.

In each plot, we use the following symbols to indicate the outcome of the ANOVA test, where p is the p-value obtained:

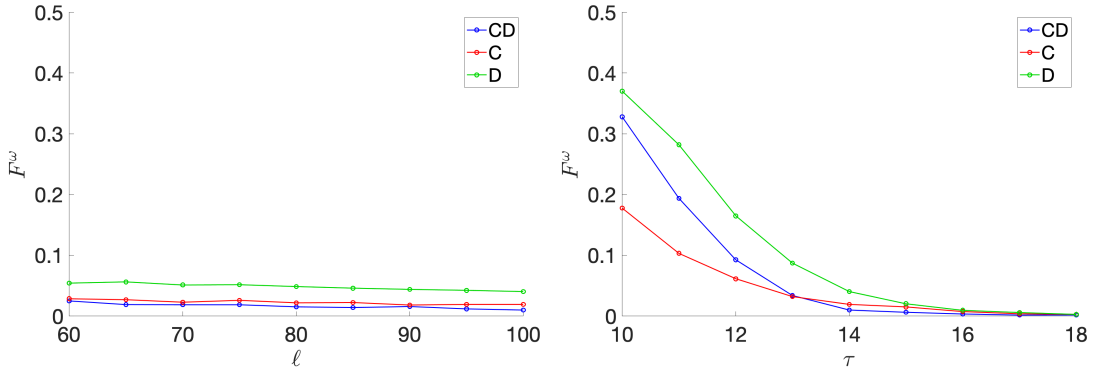
- Highly significant (***): $0 < p < 0.001$.
- Very significant (**): $0.001 < p < 0.01$.
- Significant (*): $0.01 < p < 0.05$.
- Slightly significant (.): $0.05 < p < 0.01$.
- Not significant ( ): $0.1 < p < 1$.

ANOVA tests are only one way of comparing models here, and only tell us whether *all* of the means are significantly different. Other tests such as Welch's t-test (with appropriate multi-comparison adjustments) could be implemented to check whether each pair of models have a significantly different mean.



(a) F^ω as ℓ varies ($\tau = 14$). (b) F^ω as τ varies ($\ell = 10$).

Figure 6.5: Plots of F^ω for Models CD, C and D, using 10,000 simulations per model. All other parameters are fixed. Observe that smaller values of ℓ lead to higher probabilities of mortality.

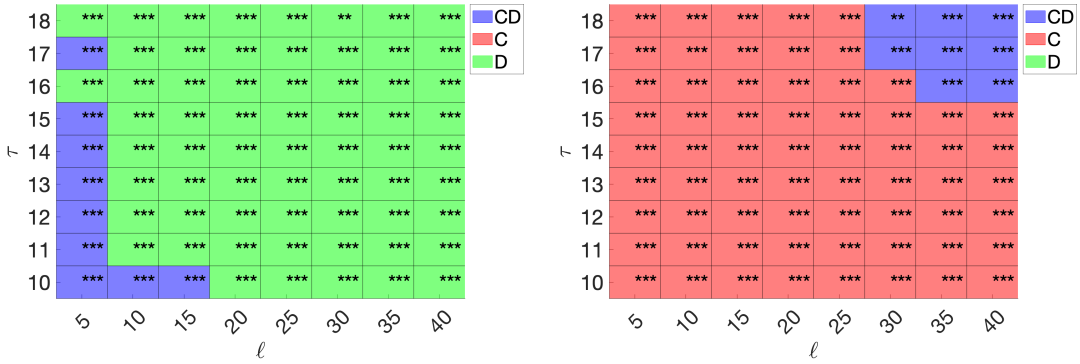


(a) F^ω as ℓ varies ($\tau = 14$). (b) F^ω as τ varies ($\ell = 100$).

Figure 6.6: Plots of F^ω for Models CD, C and D, using 10,000 simulations per model. All other parameters are fixed. Observe that in comparison to Figure 6.5, ℓ has much less of an effect on F^ω, and τ appears to have a very different effect.

in comparison to Figure 6.5a. Perhaps this is indicative that there is a 'warm up' period associated with smaller values of ℓ, but beyond $\ell \approx 60$, the process remains relatively consistent. Interestingly, Model CD appears to be the most favourable for all such values of ℓ when $\tau = 14$. In Figure 6.6b, we observe that when $\ell = 100$, τ has a very different effect on F^ω. There is a more pronounced difference between each model for $\tau = 10$, and a much steeper decay in F^ω as τ increases.

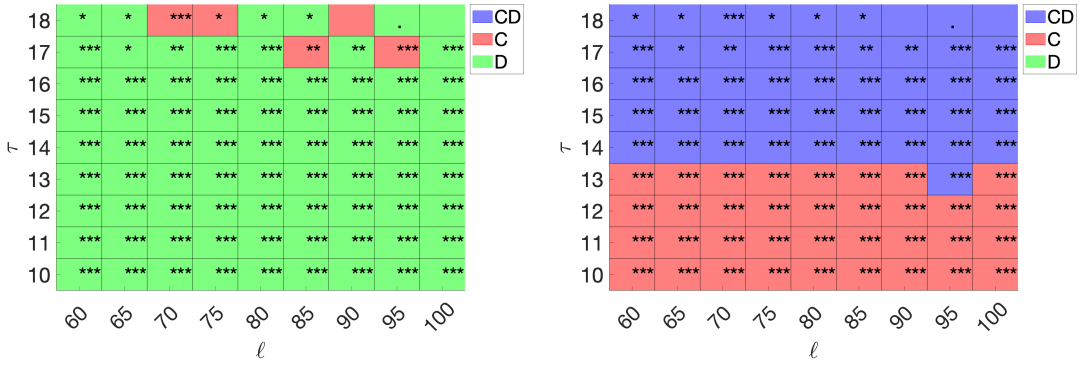
Taking a step further, we compare each model as both ℓ and τ vary, with the results displayed in Figure 6.7 (smaller ℓ) and Figure 6.8 (larger ℓ). The left-hand plots illustrate which model produced the largest value of F^ω (corresponding to the worst chance of survival), whereas the right-hand plots illustrate which model produced the smallest value of F^ω (corresponding to the best chance of survival).



(a) Largest value of F^{ω} (worst chance of survival).

(b) Smallest value of F^{ω} (best chance of survival).

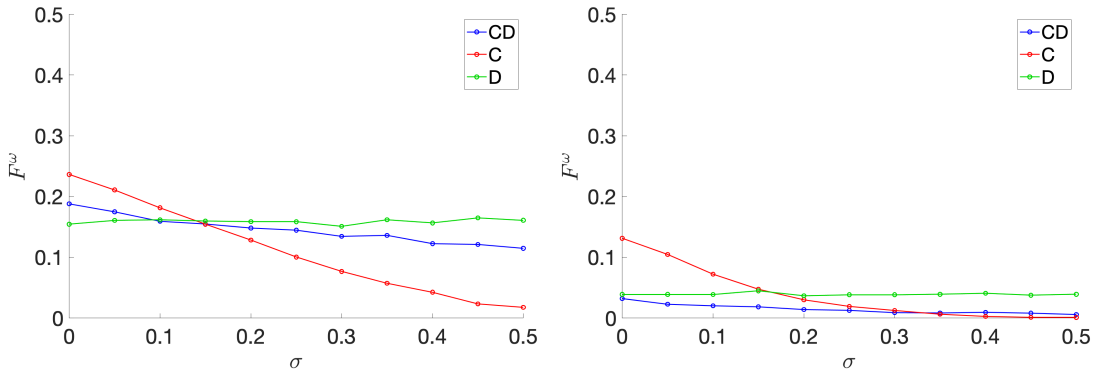
Figure 6.7: Heatmap of F^{ω} for $\ell \in \{5, 10, \ldots, 40\}$, $\tau \in \{10, 11, \ldots, 18\}$, for Models CD, C and D, using 10,000 simulations per set of parameters, per model. All other parameters are fixed. Asterisks indicate degree of statistical significance (see Comparison methods). For these parameters, Model D generally appears to have the worst chance of survival, and Model C the best chance.



(a) Largest value of F^{ω} (worst chance of survival).

(b) Smallest value of F^{ω} (best chance of survival).

Figure 6.8: Heatmap of F^{ω} for $\ell \in \{60, 65, \ldots, 100\}$, $\tau \in \{10, 11, \ldots, 18\}$, for Models CD, C and D, using 10,000 simulations per set of parameters, per model. All other parameters are fixed. Asterisks indicate degree of statistical significance (see Comparison methods). Again, we see that Model D has the worst chance of survival in most cases; however, Model CD or Model C



(a) $\ell = 10$ (b) $\ell = 100$

Figure 6.9: Plot of F^ω for Models CD, C and D, as σ varies, using 10,000 simulations per value of σ, per model. All other parameters are fixed.

For each set of parameters, the symbol in the shaded region denotes the degree of statistical significance obtained by the one-way ANOVA test (see Comparison methods for more information). In most cases, the comparison is highly significant, but we see several cases which are less significant for large values of ℓ and τ (this is likely caused by a very small probability of mortality).

We observe in Figure 6.7a that Model D appears to have the highest values of F^ω in general, with the exception of $\ell = 5$, in which Model CD produces the highest value, and for $\ell = 10, 15$ when $\tau = 10$. In Figure 6.7b, we observe that Model C produces the smallest value of F^ω in most cases, although Model CD produces the smallest value when both ℓ and τ are large.

In Figure 6.8a, we see that for larger values of ℓ, Model D once again produces the largest values of F^ω for most parameter values, with some instances of Model C producing the largest value. In Figure 6.8b, we see a very clear split between Models CD and C. For $\tau < 14$, Model C produces the smallest value of F^ω, but for $\tau \geq 14$, Model CD produces the smallest value.

Effect of σ

We would like to see what effect the parameter σ, which describes the energy advantage of genus C over genus D, has on F^ω. Figure 6.9a shows the result of varying σ in the range $[0, 0.5]$. As we might expect, F^ω decreases under Models CD and C, whereas σ has no effect on Model D.

The curve corresponding to Model C decreases, indicating that genus C becomes more preferable as σ increases, but this effect is much less pronounced in the curve for Model CD. Perhaps the presence of genus D lessens the benefit provided by genus C.

Model C appears to be the clear favourite when $\sigma > 0.15$. Overall, it is clear that σ has a great impact on the ordering of the models.

In Figure 6.9b, we repeat our process with $\ell = 100$, and observe very similar trends. The only differences between Figure 6.9 appear to be the original vertical positions of each curve.

Again, as σ increases there is a negative trend associated with Models CD and C, although Model C is no longer the clear favourite. In fact, Model CD appears to produce the smallest values of F^{ω} up until $\sigma = 0.3$, beyond which it is very similar to Model C.

Effect of horizon rates

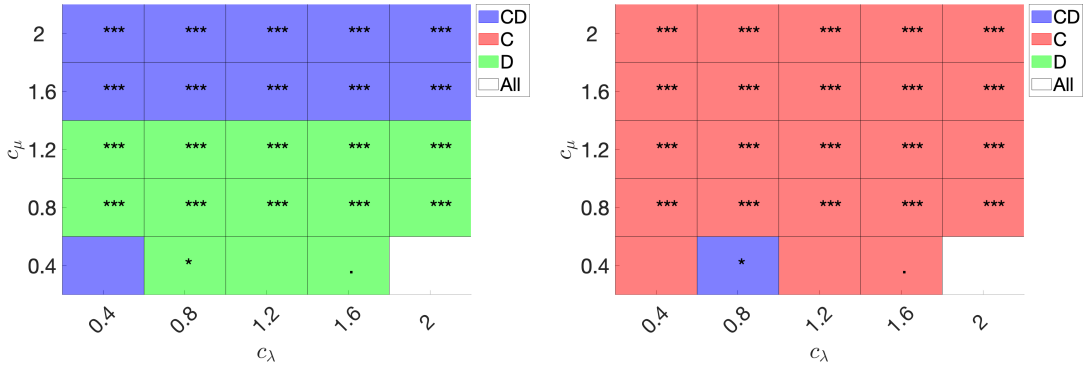
We are interested in the separate effects of λ_x and μ_x; specifically, we would like to know whether increasing one but not the other has a significant effect on what model has a higher probability of mortality. To test this, we multiply λ_c and λ_d by a factor c_λ, and similarly multiply μ_c and μ_d by a factor c_μ. Larger values of c_λ and c_μ correspond to higher rate of energy increase and decrease, respectively.

In Figure 6.10a and Figure 6.10b, we vary both c_λ and c_μ and plot a heatmap, indicating which model produced the largest and smallest value of F^{ω}, respectively. Again, the symbols in each region indicate the degree of statistical significance. The white regions indicate that all models produced the exact same value of 0 (so no simulations terminate before ω). We immediately observe that c_λ has very little effect on which model gives the worst/best chance of survival. In addition, we see that when $c_\mu = 0.4$, the results are not significant. Once again, this is likely due to the low probability of mortality associated with a small value of c_μ.

In Figure 6.10a, Model CD produces the largest value of F^{ω} for $c_\mu > 1.2$, whereas Model D produces the largest value for $c_\mu \leq 1.2$. In Figure 6.10b, we see that Model C produces the smallest value of F^{ω} for all significant tests.

We repeat the experiment for $\ell = 100$ to produce Figure 6.11. Again, c_λ does not seem to have much effect on the ordering of the models. In Figure 6.11a, we see that once again, Model CD produces the largest values of F^{ω} for $c_\mu > 1.2$, but when $c_\mu \leq 1.2$, both Models C and D produce the largest value.

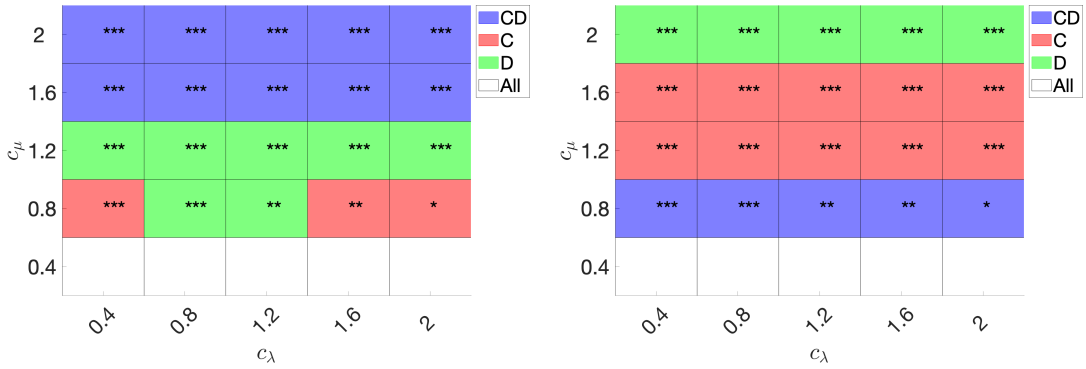
In Figure 6.11b, we see much more diversity in models which produce the smallest value of F^{ω}. For larger c_μ, Model D appears to be the most favourable; for smaller c_μ, Model CD is the most favourable; for c_μ in between, Model C is most favourable. For $c_\mu = 0.4$, we see that all models produce the same value of 0 (white regions), which can be attributed to a very small probability of mortality.



(a) Largest value of F^ω (worst chance of survival).

(b) Smallest value of F^ω (best chance of survival).

Figure 6.10: Heatmap of F^ω, for c_λ, $c_\mu \in \{0.4, \ldots, 2\}$, using 10,000 simulations per set of parameters, per model, for $\ell = 10$. All other parameters are fixed. White regions indicate that all models produced the exact same value of 0. Asterisks indicate degree of statistical significance (see Comparison methods). Observe that there is very little variation due to c_λ.



(a) Largest value of F^ω (worst chance of survival).

(b) Smallest value of F^ω (best chance of survival).

Figure 6.11: Heatmap of F^ω, for c_λ, $c_\mu \in \{0.4, \ldots, 2\}$, using 10,000 simulations per set of parameters, per model, for $\ell = 100$. All other parameters are fixed. White regions indicate that all models produced the exact same value of 0. Asterisks indicate degree of statistical significance (see Comparison methods). Once again, we observe very little variation due to c_λ. In comparison to Figure 6.10, there is much more diversity among models which produce the best chance of survival.

Effect of ζ_C and ζ_D

The thresholds ζ_C and ζ_D are the final parameters for which we would like to explore their effects. In the energy model, these thresholds play a critical role in the process of mortality, and furthermore, they drive a wedge between Models CD, C and D. Changing ζ_C does not affect Model D, but should affect Model CD and C (and vice-versa for ζ_D affecting Models CD and D, but not C). Thus, we expect that the 'ranking' of models should change with the thresholds.

Once again, we test which model produces the highest/lowest value of F^ω, for $\ell = 10$ and $\ell = 100$, shown in Figure 6.12 and Figure 6.13, respectively. Since we enforce $\zeta_C \leq \zeta_D$, we do not test the lower triangular region of the parameter space, so these regions are greyed out. Again, the statistical significance of each test is indicated by the symbols in each region. Most results are highly significant, barring two cases (these occur near a switch in model rankings).

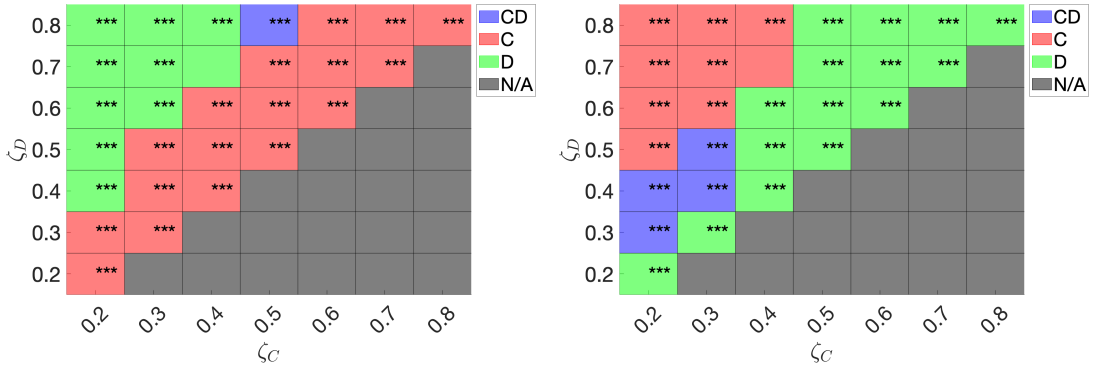
There is much to see here. We tabulate our results for $\ell = 10$ and $\ell = 100$ in Table 6.1 and Table 6.2, respectively. Although somewhat derivative, we simplify our findings by summarising the model rankings according to the relative magnitudes of ζ_C and ζ_D: close ($\zeta_C \approx \zeta_D$), far ($\zeta_C < \zeta_D$) and very far ($\zeta_C \ll \zeta_D$).

	$\zeta_C \approx \zeta_D$	$\zeta_C \ll \zeta_D$
Model CD	Best for small values, otherwise between C and D	Between C and D (except worst when $(\zeta_C, \zeta_D) = (0.5, 0.8)$)
Model C	Worst	Best
Model D	Best	Worst

Table 6.1: Comparison of Models CD, C and D for $\ell = 10$, when critical thresholds are close ($\zeta_C \approx \zeta_D$) and when they are more separated ($\zeta_C \ll \zeta_D$).

	$\zeta_C \approx \zeta_D$	$\zeta_C < \zeta_D$	$\zeta_C \ll \zeta_D$
Model CD	Between C and D	Best	Between C and D
Model C	Worst	Worst for smaller difference	Best
Model D	Best	Worst for bigger difference	Worst

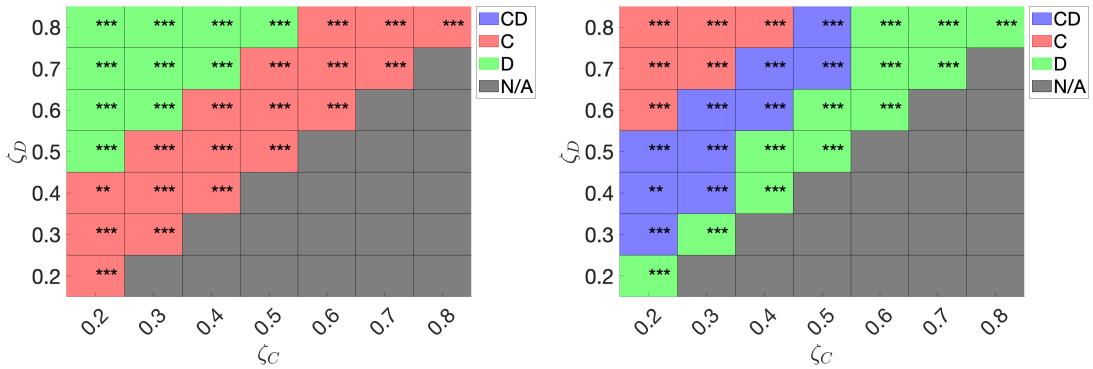
Table 6.2: Comparison of Models CD, C and D for $\ell = 100$, when critical thresholds are close ($\zeta_C \approx \zeta_D$), slightly separated ($\zeta_C < \zeta_D$) and when they are more separated ($\zeta_C \ll \zeta_D$).



(a) Largest value of F^{ω} (worst chance of survival).

(b) Smallest value of F^{ω} (best chance of survival).

Figure 6.12: Heatmap of F^{ω}, for $\zeta_C, \zeta_D \in \{0.2, \ldots, 0.8\}$, $\ell = 10$, using 10,000 simulations per set of parameters, per model. All other parameters are fixed. Dark grey rectangles indicate that $\zeta_C > \zeta_D$ (we impose that $\zeta_C \leq \zeta_D$, so these values are excluded). Asterisks indicate degree of statistical significance (see Comparison methods).



(a) Largest value of F^{ω} (worst chance of survival).

(b) Smallest value of F^{ω} (best chance of survival).

Figure 6.13: Heatmap of F^{ω}, for $\zeta_C, \zeta_D \in \{0.2, \ldots, 0.8\}$, $\ell = 100$, using 10,000 simulations per set of parameters, per model. All other parameters are fixed. Dark grey rectangles indicate that $\zeta_C > \zeta_D$ (we impose that $\zeta_C \leq \zeta_D$, so these values are excluded). Asterisks indicate degree of statistical significance (see Comparison methods).

6.5.4 Discussion

Analysis of results

There are clearly many intertwined relationships between parameters in the energy model. Identifying and visualising these relationships is quite difficult, but by using heatmaps we can glean some useful information.

One of the foremost observations is that our results are less significant when the probability of mortality is small. This is influenced by several parameters such as τ and c_μ, indicating that Models CD, C and D produce very similar values of F^ω, and the distributions of F^ω overlap. Although this is a relatively rare occurrence, we could obtain more significant results for these parameter values by increasing the number of simulations (and thus decreasing the variance of the distributions of F^ω).

The parameter ℓ is certainly influential in the model rankings. We observe very different results for $\ell = 10$ than for $\ell = 100$ in most cases. Notably, Model CD appears to be more favourable when ℓ is large. Considering that a larger value of ℓ better approximates a continuous energy buffer, this is good news for Model CD.

The critical thresholds ζ_C and ζ_D are also very influential, although their effects are not immediately discernible. We can, however, glean several important facts from Figure 6.12 and Figure 6.13:

- There is a distinct correlation between the best/worst model and the relative magnitudes between ζ_C and ζ_D.
- When ζ_C and ζ_D are very close, Model D is most preferable and Model C is least preferable.
- When ζ_C and ζ_D are far apart, Model C is most preferable and Model D is least preferable.
- When ζ_C and ζ_D are relatively close, Model CD is most preferable, although this appears to vary with ℓ.

The reason ζ_C and ζ_D have such a great effect on the model rankings is that they are the main factor in determining which model (out of Model 1, Model 2 and Model 3) is active.

When ζ_C and ζ_D are very close, Model 2 is seldom used, and the energy model is a lot like the basic model, since we essentially only have a single threshold. The main difference between the basic model and the energy model is that in the energy model, genus C gains an advantage via σ, but this does not appear to have

much effect if $\zeta_C \approx \zeta_D$. In the basic model, genus D is always more preferable, since it has a lower bleaching rate and faster growth rate. Thus, it should not be a surprise that Model D is favourable when $\zeta_C \approx \zeta_D$.

It also makes sense that when ζ_C and ζ_D are very far apart, Model C is more favourable. In this scenario, Model 2 is used the most frequently: under Model C, Model 2 is equivalent to Model 3, and under Model D, Model 2 is equivalent to Model 1. Therefore, Model 3 is used more frequently in Model C and Model 1 is used more frequently in Model D. Thus, genus C is mostly advantageous, and genus D is mostly disadvantageous.

A very interesting result is that in between these two regions of parameter space, where ζ_C and ζ_D are reasonably far apart, Model CD produces the lowest (best) value of F^ω. In this scenario, there is a roughly even distribution between whether Model 1, Model 2 or Model 3 is active; this is somewhat of a middle-ground between the two extremes.

There are several reasons why Model CD is the most favourable in this region of parameter space. The primary difference between Model CD and Models C and D is Model 2, which allows both increasing and decreasing energy under Model CD. When ζ_C and ζ_D are in a 'sweet spot', Model CD gains an advantage over Models C and D. The most likely reasons for this advantage are:

- Model CD has an advantage over Model C under ζ_C, since the algal density can re-grow at twice the rate if genus D is dominant.

- Model CD has an advantage over Model D under ζ_D, since genus C can become dominant, preventing further energy loss.

It seems that in the sweet spot, these advantages work best for Model CD, and as a result, Model CD produces the lowest value of F^ω. Although somewhat unintuitive, perhaps the best explanation is that in this situation, Model CD is the best of both worlds, and inherits the positive attributes of Models C and D.

Interestingly, we observe a very different result in Figure 6.9 and Figure 6.9b. The parameter σ is a key differentiator between each model, since it determines the advantage of genus C over genus D. We observe that larger values of σ favour Model CD and Model C, but this benefit is unbalanced. Model C appears to benefit greatly from larger values of σ, whereas Model CD only benefits slightly, indicating that the presence of genus D is detrimental to Model CD. This implies that Model CD inherits the *worst* parts of Models C and D, which is the complete opposite to our previous conclusion!

Modelling review

The energy model is the culmination of several key concepts we have seen in this thesis, providing a complex modelling framework for coral-algal symbiosis. The model is intricate, but as we have seen, the ability to track the energy process explicitly comes at the cost of model complexity, and the model relies heavily on its parameters.

A major difficulty experienced in this analysis is derived from the relatively large number of parameters specifying the energy model. In this section, we have explored the effects of ℓ, τ, λ_d, μ_d, λ_c, μ_c, ζ_C and ζ_D, but the canonical parameters ($\mathbf{T}$, $\mathbf{C}$, etc.) could also be worth investigating. There may be more complicated dynamics between different parameter combinations, but visualising the effects of all parameters at once is a challenging (if not impossible) task.

Our method of answering the research questions from Section 6.4 was to create single-genus models and compare them against the full energy model (Model CD) using tail probabilities. This method is basic but effective, since only a single probability is required per model. Of course, more sophisticated methods could be used, and whether it is valid to compare single-genus models in such a way may require further investigation.

It is necessary to disclaim that although the energy model attempts to be as realistic as possible, it is still just a model. We should be cognisant of the difficulties in selecting an appropriate set of parameters, and that the parameters chosen in this thesis do not necessarily reflect reality. Parameters such as ζ_C and ζ_D are particularly difficult to choose values for, due to their abstract nature.

To accurately represent a coral colony on the GBR, more care should be taken in choosing parameters. This model does not provide any particularly ground-breaking answers about coral bleaching; however, with the framework already in place, it is certainly possible to use more appropriate parameters to learn more about the process of coral bleaching.

A lesser objective of the energy model is to demonstrate the power of more complex models in coral research. The lack of analytic expression for the LST of the TTM is perhaps unconvincing in this regard, but even as it stands, the energy model is a mathematically founded framework for efficiently simulating coral bleaching in a random environment. Moreover, we have established analytic expressions for the LST of the TTM in the simpler models explored in Chapter 3 and Chapter 5, and it may be worthwhile using these models in place of the energy model since they require less parameter choices.

Coral survivability under the energy model

Under many parameter combinations, it is possible for Model CD to be the most favourable model. In general, the 'best' model appears to fluctuate, and is highly dependent on certain parameters. Namely:

- For large ℓ, the model becomes very sensitive to the choice of τ. Coral survivability tends to increase as ℓ increases, and (unsurprisingly) increases as τ increases, since τ has an inverse relationship with the rates of energy decrease.

- The scaling parameter c_μ, which determines the rate of energy decrease in all models, has a significant effect on the model rankings. This is much more noticeable for $\ell = 100$. Small values of c_μ lead to a much higher probability of survival, due to the lower rate of energy decrease in all models. The parameter c_λ has a much less noticeable effect on the mortality process.

- The scaling parameter σ, which determines the advantage of genus C over genus D, has a large impact on the model rankings and the probability of mortality, although it does not affect Model D.

- The critical thresholds determine which models (out of Model 1, Model 2 and Model 3) are active. Our original choices of $\zeta_C = 1/3$ and $\zeta_D = 2/3$ are examples of parameters which lie in the 'sweet spot' where Model CD is (sometimes) the best model. The critical thresholds have a large impact on the probability, as well as the model rankings.

Recall that a larger number of energy stages better approximates a continuous energy buffer. In this sense, a value of $\ell = 100$ is more realistic, and under this value Model CD has the highest probability of survival (at least for our original choices of $\zeta_C = 1/3$ and $\zeta_D = 2/3$).

With this results in mind, we should not forget our earlier concerns about our chosen method of comparison: by using tail probabilities to compare models, we are losing information about the survivability of the coral under each model, and our results are more prone to misinterpretation.

In addition, not all of the observed relationships have clear answers, and so we should approach these results with caution. The takeaway that Model CD *can* be the best model is perhaps not as important as the observation that the choice of best model is quite volatile, and that there are too many moving parts in the energy model to properly interpret these changes.

Revisiting research questions

Earlier in Section 6.4, we posed two research questions that we are now equipped to answer. The following respectively address Q1 and Q2, summarising the above analysis and discussion in the context of these questions:

A1. Under the energy model, there is indeed a set of parameters for which it is beneficial for a coral host to harbour both genus C and genus D, as opposed to a single genus only; however, the parameters which lead to this conclusion may not be representative of an actual coral ecosystem.

A2. Whether or not it is better for a coral host to harbour both genus C and genus D is certainly sensitive to the parameterisation of the energy model. The model is particularly sensitive to the thresholds ζ_C and ζ_D, as well as the parameters τ and ℓ. A more significant change may be experienced when changing multiple parameters at once. Although we attempted to explore a diverse parameter space, whether we investigated 'reasonable' perturbations of each parameter is subjective.

Computational improvements

The simulation algorithm used here could certainly be improved. For example, the algorithm uses the entire generator $\mathbf{T}$ to calculate the next phase, but this could be constructed 'on-the-fly' to save memory. In addition, parallelisation can speed up the simulation algorithm significantly: each simulation is self-contained, so it is possible to do many simulations simultaneously.

While collecting results, certain parameter combinations produce a very small probability of mortality, which may produce erroneous comparison results. To combat this, we use ANOVA tests to see whether the comparison is statistically significant, but there are other ways of dealing with this problem. Performing more simulations would be ideal, but this approach is computationally expensive. A better alternative would be a variance reduction method such as importance sampling, which could circumvent the issue in the first place by increasing the precision of our comparisons without increasing the number of simulations required.

As a final note on computations, it seems appropriate to revisit the issues associated with LST inversion. One of the key strengths of stochastic fluid models is their ability to obtain the distributions of time-dependent quantities via LST inversion, but as we have discussed, this process can be prohibitively slow.

> *Alas, poor LST! I inverted it, numerically,*
> *a quantity of infinite slowness, of most excellent fancy.*

6.5.5 Closing remarks

Thus ends our modelling journey which began with a simple model and progressed to a bounded level-dependent augmented Markov-modulated fluid flow model for coral-algal symbiosis. The analysis of such a model may either persuade or dissuade one that complex mathematical models can be powerful and useful. The difficulties arising from choosing an appropriate set of parameters should be clear, but hopefully the versatility of the model has shone through this. Regardless, our work is another step towards alleviating the quarrel in the coral.

Chapter 7

Conclusion and Future Work

7.1 Summary

Coral-algal symbiosis is an intricate biological process with far-reaching consequences in our world. In particular, the phenomenon of coral bleaching is a threat to coral reefs worldwide, including Australia's Great Barrier Reef (GBR). In this thesis, we attempted to understand the mechanisms of coral-algal symbiosis on the GBR with stochastic fluid models. We constructed a series of increasingly sophisticated models for this important relationship, and explored various offshoots of stochastic fluid models along the way.

We thoroughly investigated the process of Laplace-Stieltjes transform (LST) inversion, and how this process is affected by the complexity of our performance measure of choice, the *time to mortality* (TTM). A particularly interesting result was the importance of $\boldsymbol{\xi}(t)$, a pre-inverted LST used in the inversion of the LST of the TTM.

We explored two methods for LST inversion, the popular Euler method [1] and the recently developed CME method [48, 49, 51], which proved to be highly accurate, robust and the most effective at inverting the TTM in most cases.

To help inverting the LST of the TTM, we employed a novel time-shifting approach to circumvent issues introduced by the Euler method. This method can be used for any LST inversion method, but did not prove extremely effective, due to inaccuracies at the point of time-shifting. Nonetheless, the method could see future use in other applications.

As an improvement to the existing stochastic fluid model, we explored an extension using the concept of *Parisian ruin*. We obtained an expression for the TTM, but decided that LST inversion is an inefficient way of obtaining this distribution. Rather, we used simulation to obtain the distribution, and showed

that this model "converged" to the basic fluid model. Moreover, this concept inspired an entirely new model which uses an explicit energy process to model the mortality process in coral bleaching: the *energy model*.

The energy model uses a two-way horizon to keep track of a secondary quantity related to the existing fluid level; in our case, this is the energy buffer of the coral, which changes at a rate depending on the dominant algal genus. Although we could define the LST of the TTM, the final component proved difficult to obtain mathematically, and so we used simulation to obtain the distribution of the TTM. With simulation, we used tail probabilities to compare single-genus models to the full energy model and obtain results about the advantages and disadvantages of the two most prevalent algal genera on the GBR. Although our model produces different answers to this question based on the chosen set of model parameters, it would seem that hosting both species can indeed benefit the coral host.

Although we did not obtain too many meaningful results about coral-algal symbiosis on the GBR, the real success was the models we made along the way, which have paved the way for further research in this area, and are versatile tools which may be useful in other applications.

7.2 Future work

There are several directions to proceed from the work done in this thesis. On one hand, there are improvements to be made to the models specified here; on the other hand, these models can be simulated, and used for other purposes.

Some suggested directions are:

- To improve the existing model by choosing better parameters. As stated in Chapter 3, the canonical parameters are primarily chosen as to bring some semblance of realism to the model: they are not necessarily—and are unlikely to be—representative of reality. Fitting parameters to real data would allow the user to actually use our results meaningfully, or obtain their own results via our methods.

- To improve the biological realism of the stochastic fluid models presented here by enhancing the existing models. For example:

  - Adding more complex genus dynamics into the phase process. Our models only consider the dominant algal genus, but the phase process could account for the dominant genus as well as whether each genus has high or low density, etc.

  - Approximating a logistic growth rate. Constant algal growth rates are somewhat unrealistic, since in reality, algal populations adhere to logistic growth. It is possible to implement this by developing a level-dependent model where the growth rates are still constant, but change as the level increases. In fact, this is already possible within the energy model framework—one simply needs to change the matrices $\mathbf{C}^{(i)}$.
  - Using more general distributions in the phase process. The phase process used in the basic model is very simple, but it is possible to use more complicated distributions to model intra-genus dynamics more realistically—in fact, it is possible to use *almost any arbitrary* distribution, due to the magic of PH distributions!

- To create a more sophisticated model. Another way to improve the biological realism of our model is to create a more sophisticated version. In comparison to the above suggestions, this type of model would be transformative, rather than simply modifying the existing modelling framework. An example of such a model is the so-called *stochastic fluid-fluid model*, introduced in Bean and O'Reilly [10].

  In the fluid-fluid framework (say that five times fast), the energy process is a fluid model in its own right, depending on the existing model with symbiont density as the level. This is very similar to the energy model, but the energy process is an actual density; the energy model could be seen as a simpler, discretised version of the fluid-fluid model.

  The fluid-fluid framework is nowhere near as developed as the fluid framework, so much work would be required to allow computations in this model. Nonetheless, it would fulfil the needs of the energy model and more.

- To use the energy model for other applications. The energy model is useful in modelling coral algal-symbiosis, but it could also be used to model other processes where a fluid model is appropriate and we are interested in a second quantity which depends on the level of the fluid model.

  For example, the energy model could be used to model a risk process where the level $M(t)$ is capital, and the energy process $E(t)$ is credibility. Perhaps the capital can drop below 0, but this reduces the credibility of the account owner, who is allowed some freedom. If capital stays below 0 for too long (i.e., credibility reaches 0), the account owner goes bankrupt. If the capital returns to 0, the owner is safe from bankruptcy, but their capital must stay above 0 for some time before they fully regain their credibility.

Appendix A

Glossary

A.1 Acroynms

CDF: cumulative distribution function
CTMC: continuous-time Markov chain
DE: differential equation
GBR: Great Barrier Reef
LST: Laplace-Stieltjes transform
LT: Laplace transform
PDF: probability density function
PH: phase-type (distribution)
ROS: reactive oxygen species
SCV: squared coefficient of variation
SG: single-genus
SST: sea-surface temperature
TTM: time to mortality
UV: ultraviolet

A.2 Notation

General notation

$\Re[z]$: real part of $z \in \mathbb{C}$.
$\mathcal{L}\{f(t)\}(s)$: Laplace transform of $f(t)$.
$\mathcal{L}^*\{f(t)\}(s)$: Laplace-Stieltjes transform of $f(t)$.
$\mathbf{1}$: a column vector of ones. $\mathbf{1}_n$ denotes a vector of ones of length n.
$\mathbf{0}$: a matrix of zeros. $\mathbf{0}_{m\times n}$ denotes a matrix of size $m \times n$.
$\mathbf{I}$: identity matrix. $\mathbf{I}_n$ denotes the identity matrix of size $n \times n$.
$\mathbf{I}_+$ and $\mathbf{I}_-$: identity matrices of size $|\mathcal{S}_+|$ and $|\mathcal{S}_-|$ respectively.
$\boldsymbol{e}_i$: unit vector—a vector of zeros with a 1 in the i^{th} position. Length is implicit.
$\emptyset$: the empty set.

$\otimes$: Kronecker product. For $m \times n$ matrix $\mathbf{A}$ and $p \times q$ matrix $\mathbf{B}$ with $[\mathbf{A}]_{i,j} = a_{i,j}$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & \dots & a_{1,m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m,1}\mathbf{B} & \dots & a_{n,m}\mathbf{B} \end{bmatrix}.$$

$\oplus$: Kronecker sum. If $n = m$ and $p = q$, then

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I}_p + \mathbf{I}_n \otimes \mathbf{B}.$$

$\boldsymbol{x}^\intercal$, $\mathbf{X}^\intercal$: transposes of the vector $\boldsymbol{x}$ and matrix $\mathbf{X}$, respectively.
$\exp(\mathbf{X})$: matrix exponential of the matrix $\mathbf{X}$.

Basic fluid model notation

$M(t)$: fluid level at time t.
$\varphi(t)$: fluid phase at time t.
$\mathcal{S}$: state space of the phase process $\varphi(t)$.
$\mathbf{T}$: infinitesimal generator of the phase process $\varphi(t)$.
$\mathbf{C}$: diagonal matrix containing fluid input/output rates c_i for $i \in \mathcal{S}$.

$\mathcal{S}_+$: set of positive phases. $\mathcal{S}_+ = \{i \in \mathcal{S} : c_i > 0\}$.
$\mathcal{S}_-$: set of negative phases. $\mathcal{S}_- = \{i \in \mathcal{S} : c_i < 0\}$.
$\mathcal{S}_0$: set of zero phases. $\mathcal{S}_0 = \{i \in \mathcal{S} : c_i = 0\}$.

Appendix B

Algorithms

B.1 Numerical LST inversion

The SETUPILT function

The SETUPILT function sets up the inversion process, creating everything that will be used in the inversion and placing this information into a structure `S`. The pseudocode for this function is shown in Algorithm B.1.

The INVERTLT function

The INVERTLT function takes the output from SETUPILT and applies it to a LST that the user provides (i.e., a function which returns the LST evaluated at a specific value of s). Pseudocode for this function is shown in Algorithm B.2.

In addition to inverting scalar LSTs (such as $F(s) = \exp(-s)$), SETUPILT and INVERTLT can invert matrix LSTs, such as

$$F(s) = \begin{bmatrix} \exp(-s) & s \\ s^2 & 1/s \end{bmatrix}.$$

The dimension of the matrix output is passed to SETUPILT via the input `dim`. The ability to invert matrix LSTs will become very important when we consider fluid model LSTs, which often have matrix form.

Algorithm B.1 Pseudocode of the SETUPILT function. The file `iltcme.json` (see http://inverselaplace.org/ for more details) contains precomputed values of `beta` and `eta`.

```
function SETUPILT(Tset, dim, method, maxNodes)
    // Tset is the set of t values to invert the LST at
    // dim is the dimension of the LST output, if the output is a matrix
    // method specifies whether to use the Euler or CME method
    // maxNodes (n) is the maximum number of nodes/weights to obtain

    // Get nodes (beta) and weights (eta), both with maxNodes elements
    switch method
        case 'euler'
            // Create beta and eta arrays (see Section 2.3.2)
        case 'cme'
            // Get precomputed beta and eta arrays from iltcme.json
    end switch

    // S contains everything necessary for inversion
    S.nodes ← beta/Tset // array of points to evaluate the LST at (β_k/t)
    S.dim ← dim // dimension of LST
    S.weights ← eta // weights (η_k)
    S.Tset ← Tset // also used in inversion
    return S
end function
```

Algorithm B.2 Pseudocode of the INVERTLT function.

```
function INVERTLT(F,S)
    // F is the LST we want to (numerically) invert
    // S is the output from SETUPILT (Algorithm B.1)
    // maxNodes is not an input, but can be deduced from the size of S.nodes

    // Loop over all t values
    for i in {1, 2, ..., length(Tset)} do
        // Evaluate LST at nodes
        for j in {1, 2, ..., maxNodes} do
            Fvals(j) ← F(S.nodes(i,j))
        end for

        // Apply Abate-Whitt formula
        tvals(i) ← 1/S.Tset(i) * sum(real(S.weights * Fvals))      (†)
    end for

    return tvals // This is the inverse LST array
end function
```

B.1.1 Inverting the LST of the TTM

Obtaining the TTM can be broken into three steps: constructing a set of nodes (s values) to evaluate the LST of the TTM at, evaluating the LST at these nodes, and adding up the results in the right way to obtain the inverse transform. The overall process requires the following inputs:

- A set of model parameters, including:
  - Fluid model parameters, such as matrices $\mathbf{T}$ and $\mathbf{C}$;
  - Mortality parameters τ and ζ;
- A set of t values to evaluate $F_{\text{ttm}}(t)$ at, denoted `Tset`;
- A set of t values to evaluate $\boldsymbol{\xi}(t)$ at (required to evaluate $\int_0^\tau e^{-st}\boldsymbol{\xi}(t)\,dt$ and $\int_0^\tau \boldsymbol{\xi}(t)\,dt$), denoted `Txi`;
- The inversion method to use, denoted `method`;
- The maximum number of nodes to use in the Abate Whitt Framework, denoted `maxNodes`.

The output of the process (`ttm`) is an array of dimensions $s_- \times 1 \times n$, where s_- is the number of elements in $\mathcal{S}_-$ and n is the number of elements in `Tset`. The structure of the overall process is illustrated in Algorithm B.3, which makes use of the SETUPILT and INVERTLT functions.

The method of inversion (either Euler or CME) is used to determine the nodes (β_k) and weights (η_k) only; once the LST has been evaluated at the nodes, the results are multiplied by their corresponding weights and added together. This addition is independent of the method chosen.

Algorithm B.3 Pseudocode for inverting the LST of the TTM.

```
// Import parameters
model ← FLUIDMATRICES(rates, T, Pup, Tup, Pdown, Tdown)
Define method, maxNodes, tau, zeta, Tset, and Txi

// Create nodes and weights corresponding to points in Txi
Sxi ← SETUPILT(Txi, model, method, maxNodes)

// Define function which calculates ζ_0Ξ̃(s)
function GETXI(s, model)
    ⋮
    return XI(s)
end function

// Invert ζ_0Ξ̃(s) at points in Txi
xi ← INVERTLT(GETXI, Sxi)
// Calculate ∫_0^τ ξ(u) du via trapezoidal integration
xiprob ← TRAPZ(Txi, xi)

// Create nodes and weights corresponding to points in Tset
Sttm ← SETUPILT(Tset, model, method, maxNodes)

// Define function which calculate F̂_TTM(s)
function GETTTM(s, model, xi, Txi, xiprob)
    ⋮
    return TTM(s)
end function

// Invert F̂_TTM(s)
ttm ← INVERTLT(GETTTM, Sttm)
```

B.2 Simulation

To simulate a fluid model, we use Algorithm B.4. This function is used throughout Chapters 4, 5 and 6. The algorithm simulates a single path in a bounded fluid model (it is still possible to simulate an unbounded model with the right input parameters). To simulate something more complicated—such as the time to mortality—we can simulate multiple paths and add the results.

The inputs to the algorithm are:

- `model`: the model structure containing matrices $\mathbf{T}$, $\mathbf{C}$, etc.
- `level`: the initial level.
- `phase`: the index of the initial phase in $\mathbf{T}$.
- `dest`: the destination level. The process terminates when it reaches this level, and returns `flag=1`.
- `taboo`: the taboo level. The process terminates when it reaches this level, and returns `flag=0`.
- `b`: the boundary level. As long as the necessary matrices are contained in `model`, the process will act appropriately at this boundary. Note: the lower boundary is implicitly 0.
- `Tmax`: the maximum time allowed. If a path takes longer than `Tmax` to hit either `dest` or `taboo`, then the algorithm terminates and returns `flag=-1`.

To simulate an unbounded model, one can set `taboo` and `b` to be infinity (or negative infinity in the case of `taboo`).

Extra behaviour for horizon models

As mentioned in Chapter 5, we can use Algorithm B.4 to simulate fluid models with a horizon. To do this, we construct a new $\mathbf{T}$ matrix at (†) which includes a cemetery state. At (‡), we check whether the process has transitioned into the cemetery state, and terminate the process if so. This method is used to simulate the models from Chapter 5 and Chapter 6.

In Algorithm B.5, we present pseudocode for setting up a fluid model with a horizon, and as an example, how we construct the Erlangization model from Section 5.3.

Algorithm B.4 Pseudocode of the `returntime` function. This function simulates a path in a fluid model, outputting the time elapsed during the journey.

```
function RETURNTIME(model, level, phase, dest, taboo, b, Tmax)
    // Import model parameters
    for X in {T, C, Pup, Tup, Pdown, Tdown} do
        X ← model.X
    end for
    t ← 0 // Initialise t, which tracks the time elapsed
    // Extra step for Chapters 5 and 6                                    (†)

    // Classify the type of path occurring, using the position of the destination
    // and taboo levels, and the initial phase
    if ··· then
        // E.g. +- refers to a path which starts in an upwards state and ends in
        // a downwards state; +-b indicates the path is bounded (above)
        thecase ← one of {++, +-, -+, --, ++b, +-b, -+b, --b}
    end if

    while t < Tmax do
        // Evolve fluid
        qj = -T[phase, phase]
        sojourn = EXPRND(1/qj) // Exponential random number generation
        fluidrate = rates[phase]
        newlevel = level + fluidrate * sojourn // Candidate level

        // Check if process has hit its destination, taboo or boundary
        switch thecase
            case ++
                if newlevel > dest then
                    flag ← 1
                    level ← dest
                    t = t + (dest - level)/fluidrate;
                    break
                else if newlevel < taboo then
                    flag ← 0
                    level ← taboo
                    t = t + (taboo - level)/fluidrate;
                    break
                else
                    sendto ← newphase;
                end if
                ⋮
                // Repeat for all types of paths. If the path is bounded and hits
                // the boundary, then sendto will be set to upper or lower
        end switch
```

```
        // Update phase, level and time
        switch sendto
            case upper
                t = t + (b - level)/fluidrate;
                level ← b
                if stay at boundary then
                    // Stay at level b according to Pup
                    while at boundary do
                        t ← t + ... // Spend time at boundary
                    end while
                    update phase according to Tup[phase,:]
                else
                    // Reflect into S_
                    update phase according to Tup[phase,:]
                end if
                // Extra step for Chapters 5 and 6                    (‡)
            case lower
                t = t + (0 - level)/fluidrate;
                level ← 0
                if stay at boundary then
                    // Stay at level 0 according to Pdown
                    while at boundary do
                        t ← t + ...
                    end while
                    update phase according to Tdown[phase,:]
                else
                    // Reflect into S+
                    update phase according to Tdown[phase,:]
                end if
                // Extra step for Chapters 5 and 6                    (‡)
            case newphase
                level ← newlevel
                t ← t + sojourn
                update phase according to T[phase,:]
                // Extra step for Chapters 5 and 6                    (‡)
        end switch
    end while
    if t > Tmax then
        flag ← −1
    end if
end function
```

Algorithm B.5 Pseudocode for forming the bounded augmented model from Section 5.2. The model is specified by `kappa`, `K`, and an existing fluid model. The Erlangization model is created by specifying `kappa` and `K` as in Eq. (5.6) and Eq. (5.7), respectively.

```
function FLUIDPH(model, kappa, K)
    // Import model parameters
    for X in {T, C, Pup, Tup, Pdown, Tdown} do
        X ← model.X
    end for

    // Create new matrices
    T' ← T ⊕ K
    C' ← C ⊗ K
    ⋮
    Tdown' ← Tdown ⊕ K

    // Save new matrices into new model structure
    for Y in {T', C', Pup', Tup', Pdown', Tdown'} do
        newmodel.Y ← Y
    end for
    return newmodel
end function

// Example: creating an Erlangization model with μ = 1, ℓ = 4
model ← FLUIDMATRICES(rates, T, Pup, Tup, Pdown, Tdown)
mu ← [1, 1, 1, 1]
kappa ← [0, 0, 0, 1]
K ← LOWERDIAG(mu) - DIAG(mu) // This is the construction in Eq. (5.7)
modelPH ← FLUIDPH(model, kappa, K)
```

Algorithm B.6 Pseudocode for simulating the energy model. The parameter `tmax` is the maximum time allowed before the simulation times out. By setting `tmax` to $\omega + \varepsilon$ for some small $\varepsilon > 0$, we can cut down on simulation time.

```
// Create energy 'supermodel' structure
// This follows the level-dependent model construction in Section 6.3
supermodel ← FLUIDENERGYMODELS(...)

function SIMULATEENERGYTTM(supermodel)
    // Initialise time, level and phase
    // ph0 is the initial phase in T
    // e0 is the initial energy out of ell
    t ← 0
    lev ← b
    ph ← (ph0 - 1) × ell + e0
    // Initialise location, destination and taboo
    location ← b
    dest ← zetaD
    tab ← -Inf
    // Initialise model as Model 1
    ii ← 1
    while exitflag == 0 and t < tmax do
        // Simulate path
        [u,ph,flag,ph] ← RETURNTIME(supermodel(ii), lev, ph, dest, tab)
        t ← t+u // Update total time elapsed
        k ← k+1 // Count total number of paths

        // Define new interval based on output from RETURNTIME
        // flag indicates the success of the path in reaching its destination
        switch flag
            case 2
                // Mortality occurred
                exitflag ← 1
            case 1
                // Process hit destination (extra checks to decide which location)
                location ← zD+, zD- or zC+
            case 0
                // Process hit taboo (case where level hits ζ_C from above)
                location ← zC-
            case -1
                // Simulation timed out in RETURNTIME
                exitflag ← −1
                t ← NaN
                break
        end switch

        continued over the page ...
```

```
        // Set up next simulation
        switch location
            // Re-define dest, tab and ind depending on current location
            // dest is either zetaD, or zetaC if the level is under ζ_C
            // tab is zetaC if the level is in (ζ_C, ζ_D), or -Inf otherwise
            // ii indicates whether to use Model 1, Model 2 or Model 3
            ⋮
        end switch
    end while

    if t > tmax then
        exitflag ← −1
        t ← NaN
    end if
    return t // This is the time of mortality (NaN if the simulation timed out)
end function
```

Bibliography

[1] Abate, J., and Whitt, W. Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on computing 7*, 1 (1995), 36–43.

[2] Abate, J., and Whitt, W. A unified framework for numerically inverting Laplace transforms. *INFORMS Journal on Computing 18*, 4 (2006), 408–421.

[3] Anthony, K. R. N., Hoogenboom, M. O., Maynard, J. A., Grottoli, A. G., and Middlebrook, R. Energetics approach to predicting mortality risk from environmental stress: a case study of coral bleaching. *Functional ecology 23*, 3 (2009), 539–550.

[4] Asmussen, S., Avram, F., and Usabel, M. Erlangian approximations for finite-horizon ruin probabilities. *ASTIN Bulletin: The Journal of the IAA 32*, 2 (2002), 267–281.

[5] Baird, A. H., Bhagooli, R., Ralph, P. J., and Takahashi, S. Coral bleaching: the role of the host. *Trends in Ecology & Evolution 24*, 1 (2009), 16–20.

[6] Baird, M. E., Mongin, M., Rizwi, F., Bay, L. K., Cantin, N. E., Soja-Woźniak, M., and Skerratt, J. A mechanistic model of coral bleaching due to temperature-mediated light-driven reactive oxygen build-up in zooxanthellae. *Ecological modelling 386* (2018), 20–37.

[7] Baker, A. C. Flexibility and specificity in coral-algal symbiosis. *Annual Review of Ecology, Evolution, and Systematics* (2003), 661–689.

[8] Baker, A. C., Starger, C. J., McClanahan, T. R., and Glynn, P. W. Coral reefs: corals' adaptive response to climate change. *Nature 430*, 7001 (2004), 741.

[9] Baker, D. M., Andras, J. P., Jordán-Garza, A. G., and Fogel, M. L. Nitrate competition in a coral symbiosis varies with temperature

among Symbiodinium clades. *The ISME journal: Multidisciplinary Journal of Microbial Ecology 7*, 6 (2013), 1248.

[10] BEAN, N. G., AND O'REILLY, M. M. The stochastic fluid–fluid model: a stochastic fluid model driven by an uncountable-state process, which is a stochastic fluid model itself. *Stochastic Processes and their Applications 124*, 5 (2014), 1741–1772.

[11] BEAN, N. G., O'REILLY, M. M., AND TAYLOR, P. G. Hitting probabilities and hitting times for stochastic fluid flows. *Stochastic Processes and their Applications* (2005), 1530–1556.

[12] BEAN, N. G., O'REILLY, M. M., AND TAYLOR, P. G. Algorithms for the Laplace–Stieltjes transforms of first return times for stochastic fluid flows. *Methodology and Computing in Applied Probability 10*, 3 (2008), 381–408.

[13] BEAN, N. G., O'REILLY, M. M., AND TAYLOR, P. G. Hitting probabilities and hitting times for stochastic fluid flows: the bounded model. *Probability in the Engineering and Informational Sciences 23*, 1 (2009), 121–147.

[14] BERKELMANS, R., DE'ATH, G., KININMONTH, S., AND SKIRVING, W. J. A comparison of the 1998 and 2002 coral bleaching events on the Great Barrier Reef: spatial correlation, patterns, and predictions. *Coral Reefs 23*, 1 (2004), 74–83.

[15] BIERI, T., ONISHI, M., XIANG, T., GROSSMAN, A. R., AND PRINGLE, J. R. Relative contributions of various cellular mechanisms to loss of algae during cnidarian bleaching. *PLoS One 11*, 4 (2016), e0152693.

[16] BOULOTTE, N. M., DALTON, S. J., CARROLL, A. G., HARRISON, P. L., PUTNAM, H. M., PEPLOW, L. M., AND VAN OPPEN, M. J. H. Exploring the Symbiodinium rare biosphere provides evidence for symbiont switching in reef-building corals. *The ISME journal 10*, 11 (2016), 2693.

[17] BRANDT, M. E., AND MCMANUS, J. W. Disease incidence is related to bleaching extent in reef-building corals. *Ecology 90*, 10 (2009), 2859–2867.

[18] BUDDEMEIER, R. W., AND FAUTIN, D. G. Coral bleaching as an adaptive mechanism. *BioScience 43*, 5 (1993), 320–326.

[19] CHOW, M. H., TSANG, R. H. L., LAM, E. K. Y., AND ANG JR, P. Quantifying the degree of coral bleaching using digital photographic technique. *Journal of Experimental Marine Biology and Ecology 479* (2016), 60–68.

[20] COLOMBO-PALLOTTA, M. F., RODRÍGUEZ-ROMÁN, A., AND IGLESIAS-PRIETO, R. Calcification in bleached and unbleached Montastraea faveolata: evaluating the role of oxygen and glycerol. *Coral Reefs 29*, 4 (2010), 899–907.

[21] COMMONWEALTH OF AUSTRALIA. The Great Barrier Reef, Queensland – world heritage values. Available at http://www.environment.gov.au/heritage/places/world/gbr/values, 2019. Last accessed April 2019.

[22] CONNOLLY, S. Personal communication, 2019.

[23] CRUMP, K. S. Numerical inversion of Laplace transforms using a Fourier series approximation. *Journal of the ACM 23*, 1 (1976), 89–96.

[24] CUNNING, R., MULLER, E. B., GATES, R. D., AND NISBET, R. M. A dynamic bioenergetic model for coral-Symbiodinium symbioses and coral bleaching as an alternate stable state. *Journal of theoretical biology 431* (2017), 49–62.

[25] CUNNING, R., SILVERSTEIN, R. N., AND BAKER, A. C. Symbiont shuffling linked to differential photochemical dynamics of Symbiodinium in three Caribbean reef corals. *Coral Reefs 37*, 1 (2018), 145–152.

[26] DAVID, A., AND LARRY, S. The least variable phase type distribution is Erlang. *Stochastic Models 3*, 3 (1987), 467–473.

[27] DAVIES, P. S. Effect of daylight variations on the energy budgets of shallow-water corals. *Marine Biology 108*, 1 (1991), 137–144.

[28] DAVY, S. K., ALLEMAND, D., AND WEIS, V. M. Cell biology of cnidarian-dinoflagellate symbiosis. *Microbioly and Molecular Biology Reviews 76*, 2 (2012), 229–261.

[29] DAY, J. C. Zoning - lessons from the Great Barrier Reef marine park. *Ocean & coastal management 45*, 2-3 (2002), 139–156.

[30] DONNER, S. D., SKIRVING, W. J., LITTLE, C. M., OPPENHEIMER, M., AND HOEGH-GULDBERG, O. Global assessment of coral bleaching and required rates of adaptation under climate change. *Global Change Biology 11*, 12 (2005), 2251–2265.

[31] DOUGLAS, A. E. Coral bleaching–how and why? *Marine Pollution Bulletin 46*, 4 (2003).

[32] DUBINSKY, Z., AND JOKIEL, P. L. Ratio of energy and nutrient fluxes regulates symbiosis between zooxanthellae and corals. *Pacific Science 48*, 3 (1994), 313–324.

[33] DUNN, S. R., SCHNITZLER, C. E., AND WEIS, V. M. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proceedings of the Royal Society B: Biological Sciences 274*, 1629 (2007), 3079–3085.

[34] EYNAUD, Y., NISBET, R. M., AND MULLER, E. B. Impact of excess and harmful radiation on energy budgets in scleractinian corals. *Ecological modelling 222*, 7 (2011), 1315–1322.

[35] FAUTIN, D. G., AND BUDDEMEIER, R. W. Adaptive bleaching: a general phenomenon. In *Coelenterate Biology 2003*. Springer, 2004, pp. 459–467.

[36] GOREAU, T. F., GOREAU, N. I., AND GOREAU, T. J. Coral and Coral Reefs. *Scientific American 241*, 2 (1979), 124–136.

[37] GOULET, T. L. Most corals may not change their symbionts. *Marine Ecology Progress Series 321* (2006), 1–7.

[38] GREAT BARRIER REEF MARINE PARK AUTHORITY. Reef health. Available at http://www.gbrmpa.gov.au/the-reef/reef-health, 2016. Last accessed April 2019.

[39] GROTTOLI, A. G., RODRIGUES, L. J., AND JUAREZ, C. Lipids and stable carbon isotopes in two species of Hawaiian corals, Porites compressa and Montipora verrucosa, following a bleaching event. *Marine Biology 145*, 3 (2004), 621–631.

[40] HARRIOTT, V. J. Coral lipids and environmental stress. *Environmental monitoring and assessment 25*, 2 (1993), 131–139.

[41] HEDLEY, J., ROELFSEMA, C., CHOLLETT, I., HARBORNE, A., HERON, S., WEEKS, S., SKIRVING, W., STRONG, A., EAKIN, C., CHRISTENSEN, T., ET AL. Remote sensing of coral reefs for monitoring and management: a review. *Remote Sensing 8*, 2 (2016), 118.

[42] HELFGOTT, A., BEAN, N., CONNOLLY, S., AND BAIRD, A. A stochastic fluid model of the adaptive bleaching hypothesis on the Great Barrier Reef. Unpublished, 2009.

[43] HILL, J., AND WILKINSON, C. Methods for ecological monitoring of coral reefs. *Australian Institute of Marine Science, Townsville 117* (2004).

[44] HØEGH-GULDBERG, O. Climate change, coral bleaching and the future of the world's coral reefs. *Marine and freshwater research 50*, 8 (1999), 839–866.

[45] HØEGH-GULDBERG, O., JONES, R. J., WARD, S., AND LOH, W. K. Ecology (communication arising): is coral bleaching really adaptive? *Nature 415*, 6872 (2002), 601.

[46] HØEGH-GULDBERG, O., MUMBY, P. J., HOOTEN, A. J., STENECK, R. S., GREENFIELD, P., GOMEZ, E., HARVELL, C. D., SALE, P. F., EDWARDS, A. J., CALDEIRA, K., ET AL. Coral reefs under rapid climate change and ocean acidification. *Science 318*, 5857 (2007), 1737–1742.

[47] HØEGH-GULDBERG, O., AND SMITH, G. J. The effect of sudden changes in temperature, light and salinity on the population density and export of zooxanthellae from the reef corals Stylophora pistillata Esper and Seriatopora hystrix Dana. *Journal of Experimental Marine Biology and Ecology 129*, 3 (1989), 279–303.

[48] HORVÁTH, G., HORVÁTH, I., ALMOUSA, S. A.-D., AND TELEK, M. Numerical inverse Laplace transformation by concentrated matrix exponential distributions. In *MAM10* (Feb. 2019).

[49] HORVÁTH, G., HORVÁTH, I., ALMOUSA, S. A.-D., AND TELEK, M. Numerical inverse Laplace transformation using concentrated matrix exponential distributions. *Performance Evaluation 137* (2020), 102067.

[50] HORVÁTH, G., HORVÁTH, I., AND TELEK, M. High order concentrated matrix-exponential distributions. *Stochastic Models* (2019), 1–17.

[51] HORVÁTH, I., TALYIGÁS, Z., AND TELEK, M. An optimal inverse Laplace transform method without positive and negative overshoot - an integral based interpretation. *Electronic Notes in Theoretical Computer Science 337* (2018), 87–104.

[52] HUGHES, T. P., BAIRD, A. H., BELLWOOD, D. R., CARD, M., CONNOLLY, S. R., FOLKE, C., GROSBERG, R., HOEGH-GULDBERG, O., JACKSON, J. B. C., KLEYPAS, J., ET AL. Climate change, human impacts, and the resilience of coral reefs. *Science 301*, 5635 (2003), 929–933.

[53] JONES, R. The ecotoxicological effects of Photosystem II herbicides on corals. *Marine Pollution Bulletin 51*, 5-7 (2005), 495–506.

[54] KLUETER, A., TRAPANI, J., ARCHER, F. I., MCILROY, S. E., AND COFFROTH, M. A. Comparative growth rates of cultured marine dinoflagellates in the genus Symbiodinium and the effects of temperature and light. *PloS one 12*, 11 (2017).

[55] LAJEUNESSE, T. C., PARKINSON, J. E., GABRIELSON, P. W., JEONG, H. J., REIMER, J. D., VOOLSTRA, C. R., AND SANTOS, S. R. Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology 28*, 16 (2018), 2570–2580.

[56] LITTLE, A. F., VAN OPPEN, M. J. H., AND WILLIS, B. L. Flexibility in algal endosymbioses shapes growth in reef corals. *Science 304*, 5676 (2004), 1492–1494.

[57] LOGAN, C. A., DUNNE, J. P., EAKIN, C. M., AND DONNER, S. D. Incorporating adaptive responses into future projections of coral bleaching. *Global Change Biology 20*, 1 (2014), 125–139.

[58] LOUGH, J. M., AND VAN OPPEN, M. J. H. Introduction: coral bleaching–patterns, processes, causes and consequences. In *Coral Bleaching*. Springer, 2018, pp. 1–8.

[59] MIEOG, J. C., VAN OPPEN, M. J. H., CANTIN, N. E., STAM, W. T., AND OLSEN, J. L. Real-time PCR reveals a high incidence of Symbiodinium clade D at low levels in four scleractinian corals across the Great Barrier Reef: implications for symbiont shuffling. *Coral Reefs 26*, 3 (2007), 449–457.

[60] MSN ENCARTA. Great Barrier Reef. Available at https://www.webcitation.org/5kwqDXqi5?url=http://encarta.msn.com/encyclopedia_761575831/Great_Barrier_Reef.html, 2009. Last accessed April 2019.

[61] MULLER, E. B., KOOIJMAN, S. A. L. M., EDMUNDS, P. J., DOYLE, F. J., AND NISBET, R. M. Dynamic energy budgets in syntrophic symbiotic relationships between heterotrophic hosts and photoautotrophic symbionts. *Journal of Theoretical Biology 259*, 1 (2009), 44–57.

[62] MUSCATINE, L., AND CERNICHIARI, E. Assimilation of photosynthetic products of zooxanthellae by a reef coral. *The Biological Bulletin 137*, 3 (1969), 506–523.

[63] MUSCATINE, L., R. MCCLOSKEY, L., AND E. MARIAN, R. Estimating the daily contribution of carbon from zooxanthellae to coral animal respiration. *Limnology and oceanography 26*, 4 (1981), 601–611.

[64] OAKLEY, C. A., AND DAVY, S. K. Cell biology of coral bleaching. In *Coral Bleaching*, second edition ed. Springer, 2018.

[65] PATTON, J. S., ABRAHAM, S., AND BENSON, A. A. Lipogenesis in the intact coral Pocillopora capitata and its isolated zooxanthellae: evidence for a light-driven carbon cycle between symbiont and host. *Marine Biology 44*, 3 (1977), 235–247.

[66] PERALTA, O. *Advances of matrix–analytic methods in risk modelling*. PhD thesis, Technical University of Denmark, 2019.

[67] RODRIGUES, A. J. Convergence of Padé kernel approximants to the delayed Dirac function. *IMA Journal of Applied Mathematics 25*, 1 (1980), 17–27.

[68] ROWAN, R. Coral bleaching: thermal adaptation in reef coral symbionts. *Nature 430*, 7001 (2004), 742.

[69] SIEBECK, U. E., MARSHALL, N. J., KLÜTER, A., AND HOEGH-GULDBERG, O. Monitoring coral bleaching using a colour reference card. *Coral Reefs 25*, 3 (2006), 453–460.

[70] STANFORD, D. A., AVRAM, F., BADESCU, A. L., BREUER, L., SOARES, A. D. S., AND LATOUCHE, G. Phase-type approximations to finite-time ruin probabilities in the Sparre-Andersen and stationary renewal risk models. *ASTIN Bulletin: The Journal of the International Actuarial Association 35*, 1 (2005), 131–144.

[71] STANLEY, G., AND VAN DE SCHOOTBRUGGE, B. The evolution of the coral-algal symbiosis and coral bleaching in the geologic past. In *Coral Bleaching*. Springer, 2018, pp. 9–26.

[72] VAN HOOIDONK, R., MAYNARD, J. A., AND PLANES, S. Temporary refugia for coral reefs in a warming world. *Nature Climate Change 3*, 5 (2013), 508.

[73] WARE, J. R., FAUTIN, D. G., AND BUDDEMEIER, R. W. Patterns of coral bleaching: modeling the adaptive bleaching hypothesis. *Ecological Modelling 84*, 1-3 (1996), 199–214.

[74] WEIS, V. M. Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. *Journal of Experimental Biology 211*, 19 (2008), 3059–3066.

[75] WEISSTEIN, E. W. Delta function. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/DeltaFunction.html. Last accessed May 2019.

[76] WEISSTEIN, E. W. Laplace transform. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LaplaceTransform.html. Last accessed May 2019.

[77] WOOLDRIDGE, S., AND DONE, T. Learning to predict large-scale coral bleaching from past events: a Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. *Coral Reefs 23*, 1 (2004), 96–108.